ARCHIVE

for

RATIONAL MECHANICS

and

ANALYSIS

Edited by

C. TRUESDELL

Volume 5, Number 5

SPRINGER-VERLAG
BERLIN-GÖTTINGEN-HEIDELBERG
(Postverlagsort Berlin · 20. 7. 1960)

Mechanicam vero duplicem Veteres constituerunt: Rationalem *quae per Demonstrationes accurate procedit, & Practicam. Ad practicam spectant Artes omnes Manuales, a quibus utique* Mechanica *nomen mutuata est. Cum autem Artifices parum accurate operari soleant, fit ut* Mechanica *omnis a* Geometria *ita distinguatur, ut quicquid accuratum sit ad* Geometriam *referatur, quicquid minus accuratum ad* Mechanicam. *Attamen errores non sunt Artis sed Artificum. Qui minus accurate operatur, imperfectior est Mechanicus, & si quis accuratissime operari posset, hic foret Mechanicus omnium perfectissimus.* NEWTON

La généralité que j'embrasse, au lieu d'éblouïr nos lumieres, nous découvrira plutôt les véritables loix de la Nature dans tout leur éclat, & on y trouvera des raisons encore plus fortes, d'en admirer la beauté & la simplicité. EULER

Ceux qui aiment l'Analyse, verront avec plaisir la Méchanique en devenir une nouvelle branche ... LAGRANGE

The ARCHIVE FOR RATIONAL MECHANICS AND ANALYSIS nourishes the discipline of mechanics as a deductive, mathematical science in the classical tradition and promotes pure analysis, particularly in contexts of application. Its purpose is to give rapid and full publication to researches of exceptional moment, depth, and permanence.

Each memoir must meet a standard of rigor set by the best work in its field Contributions must consist largely in original research; on occasion, an expository paper may be invited.

English, French, German, Italian, and Latin are the languages of the Archive. Authors are urged to write clearly and well, avoiding an excessively condensed or crabbed style.

Manuscripts intended for the Archive should be submitted to an appropriate member of the Editorial Board.

The ARCHIVE FOR RATIONAL MECHANICS AND ANALYSIS appears in numbers struck off as the material reaches the press; five numbers constitute a volume. Subscriptions may be entered through any agent. The price is DM 96.— per volume.

For each paper 75 offprints are provided free of charge.

Measure-theoretic Foundations of Statistical Mechanics

ROBERT M. LEWIS

Communicated by M. KAC

Contents

Abstract

The basic formulas of classical equilibrium statistical mechanics are derived from well-known theorems in measure theory and ergodic theory. The method used is a generalization of the methods of KHINCHIN and GRAD and deals with several, in fact a "complete set", of "invariants" or "integrals of the motion". Most of the results are simple corollaries of BIRKHOFF's ergodic theorem, and since time-averages are used, the whole approach is characterized by an absence of statistical "ensembles" and probability notions. In the course of the development a "generalized temperature" is introduced, and a generalization of the second law of thermodynamics is derived. Formulas for the "microcanonical", "canonical", and "grand canonical" distributions appear as special cases of the general theory.

Introduction

The subject of this paper is the foundations of classical equilibrium statistical mechanics. The approach used is similar to that of A. I. KHINCHIN [5] and H. GRAD [2] and is characterized by an absence of statistical "ensembles"; in fact, probability notions play no role in the theory. Our method generalizes that of KHINCHIN in two directions. First, following the suggestion of GRAD, we consider more than one invariant, or integral of the motion. Second, instead of considering transformations of phase space for which Lebesgue measure is invariant, we consider measure-preserving transformations of arbitrary measure spaces.

Section 1, the mathematical theory, may be viewed as a small addition to measure theory and ergodic theory; in fact, most of the results may be considered corollaries of BIRKHOFF's ergodic theorem. In the beginning of that section,

we use the theory of differentiation of set functions to arrive at a formula which is the natural generalization of the "microcanonical distribution". This approach, which is very simple, is made possible by the introduction of the notion of a "complete set of invariant functions". In order to treat a system in interaction with its environment we introduce a precise definition of "weak interaction"; and in deriving the generalized form of the "canonical distribution" we use GRAD's definition of a "large component" rather than KHINCHIN's approach based on the central limit theorem. In our development, the exponential law of the canonical distribution appears as a natural consequence of an attempt to extend our results by the use of equivalent measures.

The motivation of the mathematical theory appears in Section 2a. There we discuss physical systems for which an initial value problem is appropriate. We consider the "state space" of such a system and introduce on this space measures which are preserved by the "solution operator". Our approach leads, in Section 2b, to a generalized notion of temperature and, in Section 2c, to a generalization of the second law of thermodynamics. In Section 2d, we derive the usual results for classical mechanics from our general theory. There (see Lemma 6) we provide a proof of certain assertions which are frequently taken for granted. Finally, in Section 2e, we derive the "grand canonical distribution" within the framework of the general theory.

As is already evident, we shall make use of the language and basic theorems of abstract measure theory. Hence we begin with a review of a few of the basic definitions:

If Γ is any space, then $(\Gamma, \mathscr{A})$ is called a *measurable space* if $\mathscr{A}$ is a *completely additive class* of subsets of Γ, *i.e.*, if $\mathscr{A}$ contains the empty set, φ, and is closed under complementation and formation of countable unions. The elements of $\mathscr{A}$ are called *measurable sets*. If $(\Gamma, \mathscr{A})$ and $(\Gamma', \mathscr{A}')$ are measurable spaces, then a single-valued function f from Γ into Γ' is a *measurable function* or *measurable transformation* if the inverse image of every set $A' \in \mathscr{A}'$ is a set $A \in \mathscr{A}$. For real-valued functions, Γ' is the real line and, unless otherwise specified, $\mathscr{A}'$ is taken to be the class of *Borel sets*, the smallest completely additive class of sets containing all the open sets. If $(\Gamma, \mathscr{A})$ is a measurable space, then $(\Gamma, \mathscr{A}, m)$ is a *measure space* if m is a *measure function*, *i.e.*, if m is a non-negative real-valued set function defined on $\mathscr{A}$ such that

$$m\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} m(A_n)$$

for all sequences $\{A_n\}$ of disjoint sets of $\mathscr{A}$ and $m(\varphi) = 0$. Every measure space is assumed to be *σ-finite*; *i.e.*, Γ is the union of a countable collection of sets of finite measure.

Let $(\Gamma_\nu, \mathscr{A}_\nu, m_\nu)$ be measure spaces for $\nu = 1, \ldots, n$. Let A_ν be a subset of Γ_ν and let γ_ν denote a point (element) of Γ_ν. The *product set* $A_1 \times \cdots \times A_n$ is the set of all ordered n-tuples $(\gamma_1, \ldots, \gamma_n)$ where $\gamma_\nu \in A_\nu$. If each A_ν is a measurable set of finite measure, then $A_1 \times \cdots \times A_n$ is called a *rectangle*. The smallest completely additive class of subsets of $\Gamma_1 \times \cdots \times \Gamma_n$ containing all rectangles $A_1 \times \cdots \times A_n$ is called the *product class* and is denoted by $\mathscr{A}_1 \times \cdots \times \mathscr{A}_n$. It can be shown

that there exists a unique measure m on $(\Gamma_1 \times \cdots \times \Gamma_n, \mathscr{A}_1 \times \cdots \times \mathscr{A}_n)$ such that for every rectangle, $m(A_1 \times \cdots \times A_n) = \prod_{\nu=1}^{n} m_\nu(A_\nu)$. m is called the *product measure* and is denoted by $m_1 \times \cdots \times m_n$. $(\Gamma_1 \times \cdots \times \Gamma_n, \mathscr{A}_1 \times \cdots \times \mathscr{A}_n, m_1 \times \cdots \times m_n)$ is a σ-finite measure space.

If $(\Gamma, \mathscr{A}, m)$ is a measure space, a measurable transformation T from Γ into itself is said to be *measure-preserving* if* $m(T^{-1}A) = m(A)$ for every $A \in \mathscr{A}$, and a real-valued measurable function $f(\gamma)$ defined on Γ is said to be an *invariant function*, or an *invariant*, if $f(T\gamma) = f(\gamma)$ a.e. (*i.e.*, "almost everywhere", that is, except on a set of m-measure zero). If T_t is a family of measure-preserving transformations depending on the parameter t, then f is an invariant function relative to the family if there exists a set C of measure zero such that for $\gamma \in -C$, $f(T_t\gamma) = f(\gamma)$ for all t. $-C$ denotes the complement of C.

1. The mathematical theory

Let $(\Gamma, \mathscr{A}, m)$ be a measure space, and let T_t be a family of transformations of Γ into itself which are measure-preserving for $t \in \mathscr{P}$, where $\mathscr{P}$ denotes the real semi-axis $0 \leqq t < \infty$. Let $\mathscr{M}$ denote the class of Lebesgue-measurable subsets of $\mathscr{P}$. We shall say that $(\Gamma, \mathscr{A}, m, T_t)$ is a *measure-preserving space* whenever $T_{s+t} = T_s T_t$ for each t and s in $\mathscr{P}$ and $T_t\gamma$ is a measurable function of t and γ; *i.e.*, $T_t\gamma$ is a measurable transformation from $(\mathscr{P} \times \Gamma, \mathscr{M} \times \mathscr{A})$ into $(\Gamma, \mathscr{A})$.

Our work in this section is based mainly on the following theorem of Birkhoff, a proof of which is presented in [*4*].

Ergodic Theorem. If $(\Gamma, \mathscr{A}, m, T_t)$ is a measure-preserving space and $f(\gamma)$ is integrable on $(\Gamma, \mathscr{A}, m)$, then the limit

(1) $$f^*(\gamma) = \lim_{L \to \infty} \frac{1}{L} \int_0^L f(T_t\gamma)\, dt$$

exists a.e. (The integral in (1) is a Lebesgue integral.) $f^*(\gamma)$ is integrable and invariant relative to the family T_t. If $m(\Gamma)$ is finite then

(2) $$\int f^*(\gamma)\, dm = \int f(\gamma)\, dm.$$

Given a measure-preserving space $(\Gamma, \mathscr{A}, m, T_t)$, a set of real-valued measurable functions $y_1, \ldots, y_k$ will be called a *complete set of invariant functions*, and the vector $y(\gamma) = \{y_1, \ldots, y_k\}$ will be called a *complete invariant vector* if each of the functions $y_j(\gamma)$ is an invariant relative to the family T_t, and if every real-valued measurable invariant function $z(\gamma)$ is a measurable function of $\{y_1, \ldots, y_k\}$ a.e. (If R^k denotes k-dimensional euclidean space, and B^k denotes the class of Borel sets in R^k, a *Borel measurable* function is one which is measurable from (R^k, B^k) to (R^1, B^1). We require that $z(\gamma) = Z[y(\gamma)]$ a.e., where Z is Borel measurable.) For convenience we shall say that $(\Gamma, \mathscr{A}, m, T_t, y)$ is a *complete space* whenever $(\Gamma, \mathscr{A}, m, T_t)$ is a measure-preserving space, and y is a complete invariant vector.

---

* $T^{-1}A$ denotes the set of all elements γ such that $T\gamma$ is an element in the set A.

If $(\Gamma, \mathscr{A}, m, T_t, y)$ is a complete space and f is integrable, then according to the Ergodic Theorem

$$f^*(\gamma) = F[y(\gamma)] \quad \text{a.e.} \tag{3}$$

Thus we can evaluate f^* if we can evaluate F. In what follows we shall show how this can be done.

Given any Borel set B in B^k, let $A = y^{-1}(B) = \{\gamma \ni y(\gamma) \in B\}$*. Then $A \in \mathscr{A}$ and for all $t \geqq 0$, T_t maps $A - C$ into itself, where C is a set of measure zero. Accordingly we may apply the Ergodic Theorem, using as the basic space not Γ, but $A - C$. It follows from (2) that if $m(A) < \infty$

$$\int_{A-C} f^* \, dm = \int_{A-C} f \, dm \tag{4}$$

and since $m(C) = 0$

$$\int_A f^* \, dm = \int_A f \, dm. \tag{5}$$

Let M be the measure induced in (R^k, B^k) by y and m, *i.e.*, for any $B \in B^k$

$$M(B) = m[y^{-1}(B)]. \tag{6}$$

It can be shown that if G is a Borel measurable function such that $G[y(\gamma)]$ is integrable on $(\Gamma, \mathscr{A}, m)$, then G is integrable on (R^k, B^k, M) and

$$\int_{y^{-1}(B)} G[y(\gamma)] \, dm = \int_B G(y) \, dM. \tag{7}$$

It follows now from (3), (5), and (7) that for any $B \in B^k$ such that $m[y^{-1}(B)] < \infty$

$$\int_B F(y) \, dM = \int_{y^{-1}(B)} F[y(\gamma)] \, dm = \int_{y^{-1}(B)} f^* \, dm = \int_{y^{-1}(B)} f \, dm. \tag{8}$$

Integrability of F follows from that of f^*.

In order to evaluate F we shall apply a generalization of the well-known theorem which asserts that the derivative of an indefinite integral is equal a.e. to the integrand. We state, without proof, the following lemma, which follows easily from the theory of differentiation of set functions with respect to "nets". A full discussion of the theory appears in [*6*].

Lemma 1. *Let B^k be the class of Borel sets in R^k and let (R^k, B^k, M) be a measure space such that M assigns finite values to finite intervals. Let F be integrable on (R^k, B^k, M). Then for almost every $x \in R^k$ there exists a sequence of half-open intervals*

$$I_j(x) = \{y \ni c^j < y - x \leqq d^j\} = \{y_1, \ldots, y_k \ni c^j_\nu < y_\nu - x_\nu \leqq d^j_\nu;\ \nu = 1, \ldots, k\}$$

such that for each $j = 1, 2, \ldots$; $c^j < 0 \leqq d^j$; $\lim\limits_{j\to\infty} (d^j - c^j) = 0$; $I_{j+1}(x) \subset I_j(x)$; and

$$\lim_{j\to\infty} \frac{\int_{I_j(x)} F \, dM}{M[I_j(x)]} = F(x). \tag{9}$$

* The expression $\{\gamma \ni y(\gamma) \in B\}$ means: The set of all elements γ such that $y(\gamma)$ is an element of the set B.

Now from (8) and (9) we have

$$F(x) = \lim_{j\to\infty} \frac{\int\limits_{y^{-1}[I_j(x)]} f\,dm}{m\{y^{-1}[I_j(x)]\}} \quad \text{a.e.} \tag{10}$$

From (3) we see that f^* has the same value for all γ for which $y(\gamma)$ has the same value. If $y^{-1}(x)$ denotes the set $\{\gamma \ni y(\gamma) = x\}$ it is natural to write $f^*[y^{-1}(x)]$ for the value of $f^*(\gamma)$ common to all γ in $y^{-1}(x)$. Summarizing our results, we have:

Theorem 1. *Let $(\Gamma, \mathscr{A}, m, T_t, y)$ be a complete space such that for every finite interval I in R^k, $m[y^{-1}(I)]$ is finite. Let f be integrable on $(\Gamma, \mathscr{A}, m)$. Then*

$$f^*(\gamma) = \lim_{L\to\infty} \frac{1}{L} \int_0^L f(T_t\gamma)\,dt$$

exists a.e. There exist sequences c^j and d^j of points in R^k which depend on x, for which $c^j < 0 \leqq d^j$, $\lim\limits_{j\to\infty} (d^j - c^j) = 0$, and

$$f^*[y^{-1}(x)] = \lim_{j\to\infty} \frac{\int\limits_{\{\gamma\ni c^j<y(\gamma)-x\leqq d^j\}} f\,dm}{m\{\gamma \ni c^j < y(\gamma) - x \leqq d^j\}} \quad \text{a.e.} \tag{11}$$

Theorem 1 will be applied later to describe the behavior of isolated physical systems. Equation (11) is a generalization of what is sometimes called the "microcanonical distribution". We turn now to the derivation of a formula appropriate to a physical system in interaction with its environment.

Let $y' = \{y_1', \ldots, y_k'\}$ and $y'' = \{y_1'', \ldots, y_k''\}$ be vector functions with values in R^k. The complete spaces $(\Gamma', \mathscr{A}', m', T_t', y')$ and $(\Gamma'', \mathscr{A}'', m'', T_t'', y'')$ will be said to be in *weak interaction* if there exists a complete space $(\Gamma, \mathscr{A}, m, T_t, y)$ where $\Gamma = \Gamma' \times \Gamma''$; $\mathscr{A} = \mathscr{A}' \times \mathscr{A}''$; $m = m' \times m''$; and $y(\gamma) = y(\gamma', \gamma'') = \{y_1, \ldots, y_k\} = y'(\gamma') + y''(\gamma'')$*. If $m[y^{-1}(I)]$ is finite for every finite interval I in R^k and if $f(\gamma')$ is integrable on $(\Gamma', \mathscr{M}', m')$, then f is integrable on $(\Gamma, \mathscr{A}, m)$ and we may apply Theorem 1. It follows that

$$f^*(\gamma) = \lim_{L\to\infty} \frac{1}{L} \int_0^L f(T_t\gamma)\,dt \tag{12}$$

exists a.e., and $f^*[y^{-1}(x)]$ is given by (11).

We should note that although we have taken f to be a function of γ' alone, f^* is a function of $\gamma = (\gamma', \gamma'')$ In (12), $f(T_t\gamma)$ means $f(\beta)$ where β is the first component of $T_t\gamma$. According to our assumptions, $dm = dm'dm''$ in (11), and we may use FUBINI's theorem to carry out the integration with respect to γ''. The result is

$$f^*[y^{-1}(x)] = \lim_{j\to\infty} \frac{\int f(\gamma')\, m''\{\gamma'' \ni c^j < y'(\gamma') + y''(\gamma'') - x \leqq d^j\}\, dm'}{m\{\gamma \ni c^j < y(\gamma) - x \leqq d^j\}}. \tag{13}$$

In the applications, the system of interest will be represented by Γ', and Γ'' will represent the environment, which is assumed to be, in some sense, "large". Before giving a precise definition of "large", let us note that whenever $(\Gamma', \mathscr{A}',$

---

* In the applications we shall see that it may be possible to satisfy the equation $y = y' + y''$ only in an approximate sense.

$m', T_t', y')$ and $(\Gamma'', \mathscr{A}'', m'', T_t'', y'')$ are complete spaces in weak interaction, it is possible to find new measures p', p'' such that $(\Gamma', \mathscr{A}', p', T_t', y')$ and $(\Gamma'', \mathscr{A}'', p'', T_t'', y'')$ are complete spaces in weak interaction and $p'(\Gamma') = p''(\Gamma'') = 1$. We shall see shortly how this can be done*.

Definition. $(\Gamma'', \mathscr{A}'', m'', T_t'', y'')$ is *large* compared to $(\Gamma', \mathscr{A}', m', T_t', y')$, or in symbols $(\Gamma'', \mathscr{A}'', m'', T_t'', y'') \gg (\Gamma', \mathscr{A}', m', T_t', y')$ if the two complete spaces are in weak interaction and the following condition is satisfied: Whenever p'', p' are measures such that $(\Gamma'', \mathscr{A}'', p'', T_t'', t'')$, $(\Gamma', \mathscr{A}', p', T_t', y')$ are complete spaces in weak interaction and $p'(\Gamma') = p''(\Gamma'') = 1$, there exists a vector λ in R^k (depending on p' and p'') such that for $c < 0 \leqq d$ and $d - c$ sufficiently small

$$g(y') = p''\{\gamma'' \ni c < y''(\gamma'') + y' - \lambda \leqq d\} = \text{const.} \tag{14}$$

for almost all values of y' (*i.e.*, except for y' in some Borel set B such that $p'\{\gamma' \ni y'(\gamma') \in B\} = 0$)**.

Now suppose $(\Gamma'', \mathscr{A}'', m'', T_t'', y'') \gg (\Gamma', \mathscr{A}', m', T_t', y')$, and p'', p' are measures with the above properties. Let $p = p' \times p''$, and let us use the abbreviated notation

$$p''[y'(\gamma')] = p''\{\gamma'' \ni c^j < y'(\gamma') + y''(\gamma'') - \lambda \leqq d^j\}, \tag{15}$$

$$p = p\{\gamma \ni c^j < y(\gamma) - \lambda \leqq d^j\}. \tag{16}$$

Then from (13)

$$f^*[y^{-1}(\lambda)] = \lim_{j \to \infty} \frac{\int f(\gamma')\, p''[y'(\gamma')]\, dp'}{p}. \tag{17}$$

If j is sufficiently large, then $d^j - c^j$ is as small as we wish and $p''[y'(\gamma')] = p''$ a.e. where p'' is a constant. Thus

$$\int f(\gamma')\, p''[y'(\gamma')]\, dp' = \int f(\gamma')\, dp' \cdot p'' = \int f(\gamma')\, dp' \cdot \int p''[y'(\gamma')]\, dp' = \int f(\gamma')\, dp' \cdot p \tag{18}$$

and

$$f^*[y^{-1}(\lambda)] = \int f(\gamma')\, dp'. \tag{19}$$

Equation (19) is a simple formula for $f^*(\gamma)$, but it holds only for those points γ such that $y(\gamma) = \lambda$. In order to obtain a simple formula for $f^*[y^{-1}(x)]$ for

* In case $m'(\Gamma') < \infty$ and $m''(\Gamma'') < \infty$ we need only set $p' = m'/m'(\Gamma')$, $p'' = m''/m''(\Gamma'')$.

** This definition is a generalized form of the definition of a "large component" introduced by GRAD (see [2], p. 469). Its physical interpretation is discussed further in the next section. GRAD's approach is certainly much simpler than that of KHINCHIN [5], who postulates that the environment consists of a large number, $n-1$, of weakly interacting systems more or less similar to the system of interest. In that case the quantities p'' and p appearing in (17) may each be interpreted as the probability that the sum of a large number of independent random variables falls within a certain interval. If we then choose λ to be the mean of the distribution of the sum, and if we further require that the distributions be continuous (*i.e.*, representable by density functions in R^k), then we can apply the central limit theorem and obtain the result (19) asymptotically for $n \to \infty$.

In the applications we shall not require that (14) hold exactly; rather we shall suppose that there is some parameter, l, which measures how "large" the environment is, and (14) will hold asymptotically for $l \to \infty$. Consequently (18), (19) and equations obtained from (19) will be viewed as asymptotic results.

values of x not equal to λ we shall introduce new measures. In this connection, the following lemma will be used:

Lemma 2. *Let T_t be a family of measure-preserving transformations with respect to $(\Gamma, \mathscr{A}, m)$. Let $\varphi(\gamma)$ be non-negative and integrable on $(\Gamma, \mathscr{A}, m)$. For each $A \in \mathscr{A}$, let $q(A) = \int_A \varphi(\gamma)\, dm$*. Then T_t is a family of measure-preserving transformations with respect to $(\Gamma, \mathscr{A}, q)$ if and only if φ is an invariant relative to the family T_t and $(\Gamma, \mathscr{A}, m)$.*

Proof. By definition $m(T_t^{-1}A) = m(A)$ for $A \in \mathscr{A}$.

1. Let φ be an invariant: $\varphi(T_t\gamma) = \varphi(\gamma)$ a.e. Then

$$q(T_t^{-1}A) = \int_{T_t^{-1}A} \varphi(\gamma)\, dm = \int_A \varphi(T_t\gamma)\, dm = \int_A \varphi(\gamma)\, dm = q(A).$$

Thus T_t preserves q.

2. Let T_t preserve q: $q(T_t^{-1}A) = q(A)$. Then

$$\int_A \varphi(\gamma)\, dm = q(A) = q(T_t^{-1}A) = \int_A \varphi(T_t\gamma)\, dm$$

for all A. Hence $\varphi(\gamma) = \varphi(T_t\gamma)$ a.e.; *i.e.*, φ is an invariant.

Now suppose $(\Gamma', \mathscr{A}', m', T_t', y')$, $(\Gamma'', \mathscr{A}'', m'', T_t'', y'')$ are complete spaces in weak interaction. On each space we define a new measure

$$q'(A') = \int_{A'} \varphi'(\gamma')\, dm', \qquad q''(A'') = \int_{A''} \varphi''(\gamma'')\, dm''. \tag{20}$$

By the lemma, $(\Gamma', \mathscr{A}', q', T_t', y')$ will be a complete space if and only if φ' is an invariant with respect to $(\Gamma', \mathscr{A}', m', T_t', y')$. Since y' is a complete invariant vector, this will be true if and only if $\varphi'(\gamma') = \Phi'[y'(\gamma')]$, where Φ' is a Borel-measurable function on R^k. Similarly $(\Gamma'', \mathscr{A}'', q'', T_t'', y'')$ will be a complete space if and only if $\varphi''(\gamma'') = \Phi''[y''(\gamma'')]$. Now let $q = q' \times q''$ be the product measure on $(\Gamma, \mathscr{A})$. For any rectangle

$$\begin{aligned} q(A' \times A'') &= q'(A')\, q''(A'') = \int_{A'} \Phi'[y']\, dm' \int_{A''} \Phi''[y'']\, dm'' \\ &= \int_{A'} \int_{A''} \Phi'[y']\, \Phi''[y'']\, dm'\, dm'' = \int_{A' \times A''} \Phi'[y']\, \Phi''[y'']\, dm. \end{aligned} \tag{21}$$

It follows that for any $A \in \mathscr{A}$, $q(A) = \int_A \Phi'[y']\ \Phi''[y'']\, dm$. Now, as before, $(\Gamma, \mathscr{A}, q, T, y)$ will be a complete space if and only if

$$\Phi'[y']\, \Phi''[y''] = \Phi[y] = \Phi[y' + y''], \tag{22}$$

where again Φ is Borel-measurable on R^k. (22) is a functional equation for the functions Φ', Φ'', Φ which has the solutions $\Phi'[y'] = c' e^{-\alpha \cdot y'}$; $\Phi''[y''] = c'' e^{-\alpha \cdot y''}$; $\Phi[y] = c' c'' e^{-\alpha \cdot y}$; where $\alpha = \alpha_1, \ldots, \alpha_k$ is an arbitrary vector in R^k, c' and c'' are arbitrary non-negative real numbers, and $\alpha \cdot y = \sum_{j=1}^{k} \alpha_j y_j$, *etc.* For convenience,

---

* Then q is a measure on $(\Gamma, \mathscr{A})$.

we shall refer to measures q and m which are related by equations of the form $q(A)=\int_A ce^{-\alpha\cdot y(\gamma)}\,dm$, $(A\in\mathscr{A})$, as *exponentially equivalent measures.*

Now, let us assume that the complete spaces $(\Gamma', \mathscr{A}', m', T_t', y')$ and $(\Gamma'', \mathscr{A}'', m'', T_t'', y'')$ are in weak interaction, where m' and m'' are not assumed to be normalizable, *i.e.*, we do not assume that $m'(\Gamma')$ and $m''(\Gamma'')$ are finite. We now consider the class of normalizable measures which are exponentially equivalent to m' and m'': Let D' and D'' be subsets of R^k consisting of all vectors α such that

$$z'(\alpha)=\int e^{-\alpha\cdot y'}\,dm' \quad \text{and} \quad z''(\alpha)=\int e^{-\alpha\cdot y''}\,dm'' \tag{23}$$

are, respectively, finite. We assume that $D''\supset D'$ and that D' is a non-empty set. For each α in D' set

$$z(\alpha)=z'(\alpha)\,z''(\alpha). \tag{24}$$

By FUBINI's theorem

$$z(\alpha)=\int e^{-\alpha\cdot(y'+y'')}\,dm'\,dm''=\int e^{-\alpha\cdot y}\,dm. \tag{25}$$

Now for $A'\in\mathscr{A}'$ and $A''\in\mathscr{A}''$ set

$$p'(A')=[z'(\alpha)]^{-1}\int_{A'} e^{-\alpha\cdot y'}\,dm'; \quad p''(A'')=[z''(\alpha)]^{-1}\int_{A''} e^{-\alpha\cdot y''}\,dm''. \tag{26}$$

Then if $p=p'\times p''$,

$$p(A)=[z(\alpha)]^{-1}\int_A e^{-\alpha\cdot y}\,dm \tag{27}$$

for every set $A\in\mathscr{A}$. Of course, the measures p', p'', p depend parametrically on α.

Now $p'(\Gamma')=p''(\Gamma'')=1$, and we may apply (19). We summarize our results in the following theorem.

Theorem 2. *Let $(\Gamma'', \mathscr{A}'', m'', T_t'', y'')\gg(\Gamma', \mathscr{A}', m', T_t', y')$. Let D', D'', and $z'(\alpha)$ be defined as above and let $D''\supset D'$ where D' is a non-empty set. Let $f(\gamma')\,e^{-\alpha\cdot y'(\gamma')}$ be integrable on $(\Gamma', \mathscr{A}', m')$ for each α in D'. Then*

$$f^*(\gamma)=\lim_{L\to\infty}\frac{1}{L}\int_0^L f(T_t\gamma)\,dt \tag{28}$$

exists a.e. with respect to $m'\times m''$, and for each α in D' there exists a vector λ_α in R^k such that

$$f^*[y^{-1}(\lambda_\alpha)]=\bar{f}(\alpha)=[z'(\alpha)]^{-1}\int f(\gamma')\,e^{-\alpha\cdot y'(\gamma')}\,dm'. \tag{29}$$

Theorem 2, and in particular formula (29), is our main mathematical result and will form the basis for the physical applications. However, before proceeding to these applications we shall take a closer look at some of the mathematical concepts, and derive a corollary to Theorem 2. We begin with some definitions.

Let (R^k, B^k, M) be a measure space. B^k denotes the class of Borel sets in R^k. Let E be any open half-space in R^k such that $M(E)=0$. Let F be the union of all such open half-spaces of measure zero. F is open, and its complement H

is closed. Since H is the intersection of all closed half-spaces of "full measure" M, H will be called the *closed convex hull of* M. Let $\tilde{H}$ be the complement of the closure of F. Then $\tilde{H}$ is the largest open subset of H. $\tilde{H}$ will be called the *open convex hull of* M.

Lemma 3. *If* $M(R^k) < \infty$ *and*

$$\bar{x} = \frac{\int x \, dM}{M(R^k)},$$

then $\bar{x}$ *is a point in the closed convex hull* H *of* M. *(*x *denotes a point* $\{x_1, \dots, x_k\}$ *in* R^k.*)*

Proof. Suppose $\bar{x} \notin H$. Then $\bar{x} \in F$; hence $\bar{x} \in E$ where E is an open half-space of measure zero. By a rotation of coordinates we can assume that E is the open half-space $a < x_1$ where a is a real constant. Thus $a < \bar{x}_1$. But

$$(30) \qquad \int x_1 \, dM = \int_{x_1 \leq a} x_1 \, dM \leqq a \, M(R^k).$$

Hence $\bar{x}_1 \leqq a$. This is a contradiction.

We shall make use of this lemma a little later, but now we shall use the above definitions to examine more closely the notion of a complete set of invariants. If $(\Gamma, \mathscr{A}, m, T_t, y)$ is a complete space, the invariant functions $y_1, \dots, y_k$ in a sense span the space of all invariant functions. The set $y_1, \dots, y_k$ can always be enlarged by adding invariant functions and remains a complete set. But under what conditions can we reduce the number of functions and still have a complete set? The following lemma is an answer to this question.

Lemma 4. *Let* $(\Gamma, \mathscr{A}, m, T_t, y)$ *be a complete space, where* $y = \{y_1, \dots, y_k\}$. *Let* M *be the measure induced in* (R^k, B^k) *by* y *and* m, *i.e.,* $M(B) = m[y^{-1}(B)]$ *for* $B \in B^k$, *and let* $\tilde{H}$ *be the open convex hull of* M. *If* $\tilde{H}$ *is empty, then there exists a complete set of invariants* $\{w_1, \dots, w_{k-1}\}$.

Proof. If $\tilde{H}$ is empty, it is easy to see that there exists a hyperplane $L \subset R^k$ of "full measure"; *i.e.*, there exists a set L of the form

$$(31) \qquad L = \left\{ y \ni \sum_{i=1}^{k} a_i (y_i - \mu_i) = 0 \right\}$$

such that $M(-L) = 0$. The a_i and u_i are real constants and at least one a_i is non-zero. Hence we can define a non-singular matrix (a_{ij}); $i, j = 1, \dots, k$; where $a_{ik} = a_i$; $i = 1, \dots, k$. Let us introduce the linear transformation

$$(32) \qquad w_\nu = \sum_{i=1}^{k} a_{i\nu} (y_i - \mu_i); \qquad \nu = 1, \dots, k.$$

The w_ν are invariants. Now $y^{-1}(L) = \{\gamma \ni w_k(\gamma) = 0\}$; hence $y^{-1}(-L) = \{\gamma \ni w_k(\gamma) \neq 0\}$ and $m[y^{-1}(-L)] = M(-L) = 0$. Thus $w_k(\gamma) = 0$ a.e. If $z(\gamma)$ is any invariant, then $z(\gamma) = Z[y(\gamma)] = G[w_1, \dots, w_k]$ a.e., where Z and G are Borel-measurable on R^k. Thus $z(\gamma) = G[w_1, \dots, w_{k-1}, 0] = J[w_1, \dots, w_{k-1}]$ a.e., where J is Borel-measurable on R^{k-1}. Thus $w_1, \dots, w_{k-1}$ is a complete set of invariants.

We see from this lemma that if $\tilde{H}$ is empty, we can always obtain a smaller set of invariants which is complete. If the new $\tilde{H}$ is empty, we can repeat the process, and so on, until eventually we reach a complete set of invariants for which $\tilde{H}$ is non-empty, except in the case where the single function $y(\gamma)$ is a complete set of invariants and the corresponding $\tilde{H}$ is empty. In this case the closed convex hull H consists of a single point c in R^1 and $y=c$ a.e. A measure-preserving space will be called *indecomposable** if the only invariant functions are those which are constant a.e. Whenever $\tilde{H}$ is non-empty, the complete set of invariants $y_1, \ldots, y_k$ will be said to be non-singular. Thus we have

Lemma 5. *Let $(\Gamma, \mathscr{A}, m, T_t, \{y_1, \ldots, y_k\})$ be a complete space. Then either $(\Gamma, \mathscr{A}, m, T_t)$ is indecomposable, or there exists a non-singular complete set of invariants $\{z_1, \ldots, z_j\}$, where $j \leqq k$.*

Let us consider, for a moment, the indecomposable case: $(\Gamma, \mathscr{A}, m, T_t, y)$ is a complete space, and $y=$const. a.e. If $m(\Gamma)<\infty$, we can apply the Ergodic Theorem immediately, and from (2) we see that

$$f^*(\gamma) = \lim_{L\to\infty} \frac{1}{L} \int_0^L f(T_t\gamma)\, dt = [m(\Gamma)]^{-1} \int f(\gamma)\, dm = \text{const.} \quad \text{a.e.} \tag{32}$$

If the complete spaces $(\Gamma', \mathscr{A}', m', T_t', y')$ and $(\Gamma'', \mathscr{A}'', m'', T_t'', y'')$ are in weak interaction, and both y' and y'' are constant a.e., then $y=y'+y''$ is a constant a.e. If $m'(\Gamma')$ and $m''(\Gamma'')$ are finite, then $m(\Gamma)=m'\times m''(\Gamma'\times\Gamma'')$ is finite, and if $f(\gamma')$ is integrable on $(\Gamma', \mathscr{A}', m')$, then

$$f^*(\gamma) = \lim_{L\to\infty} \frac{1}{L} \int_0^L f(T_t\gamma)\, dt = [m(\Gamma)]^{-1} \int f(\gamma')\, dm = [m'(\Gamma')]^{-1} \int f(\gamma')\, dm'. \tag{33}$$

We see from (32) and (33) that in the indecomposable case the calculations are essentially trivial. We shall not make use of these results in the applications.

The importance of Lemma 5 is that *except for the indecomposable case we may always assume that a complete space has a non-singular complete set of invariants.* We shall make use of this fact in what follows.

Let us return now to Theorem 2. The results (28) and (29) hold, of course, when $f(\gamma')$ is the vector function $y'(\gamma')$. Thus

$$y^{*\prime}[y^{-1}(\lambda_\alpha)] = \bar{y}'(\alpha) = [z'(\alpha)]^{-1} \int y'(\gamma')\, e^{-\alpha\cdot y'(\gamma')}\, dm' = \int y'(\gamma')\, dp', \tag{34}$$

where p' is defined by (26). Now let M' be the measure induced in (R^k, B^k) by m' and y'. Similarly let P' be the measure induced by p' and y'. Since M' and P' have the same sets of measure zero they have the same open and closed convex hulls $\tilde{H}$ and H. From (34)

$$\bar{y}'(\alpha) = \int_{R^k} x\, dP', \tag{35}$$

and $P'(R^k)=1$. It follows from Lemma 3 that $\bar{y}'(\alpha)$ lies in H. Thus $\bar{y}'(\alpha)$ is a single-valued function from D' into H. We shall show that this function has

* This is in agreement with the usual usage. (See *e.g.* [4], p. 25.) The terms "ergodic" and "metrically transitive" are also used.

a single-valued inverse on $\widetilde{H}$, provided D' is an open connected set. The following corollary is a generalization of a theorem of GRAD [2], which in turn is a generalization of a theorem of KHINCHIN [5].

Corollary to the Theorem 2. *Let $\widetilde{H}$ be the open convex hull of the measure M' induced in (R^k, B^k) by m' and y'. Let D' be an open connected set. Then the function $\bar{y}'(\alpha)$ defined by* (34) *has a single-valued inverse on $\widetilde{H}$, i.e., to any $c=\{c_1, \ldots, c_k\}$ in $\widetilde{H}$, there exists a unique α in D' such that $\bar{y}'(\alpha)=c$. Thus for y' in $\widetilde{H}$, $\bar{f}=\bar{f}[\alpha(\bar{y}')]$. $\bar{f}(\alpha)$ is given explicitly by* (29), *and $\alpha(\bar{y}')$ is defined implicitly by* (34)★.

Proof. If v is any non-zero vector in R^k, the equation $v \cdot (c-y')=0$ defines a hyperplane through c and perpendicular to v. Both open half-spaces bounded by this hyperplane must be of positive M'-measure; for if one were not, then every neighborhood of c would intersect F, the union of all open half-spaces of measure zero, and c would be in the closure of F, whereas it is in the complementary set. Now let

$$\psi(\alpha)=e^{\alpha\cdot c}z'(\alpha)=\int e^{\alpha\cdot(c-y')}\,dm', \tag{36}$$

$$\psi_{ij}=\frac{\partial^2\psi}{\partial\alpha_i\,\partial\alpha_j}=\int [c_i-y'_i]\,[c_j-y'_j]\,e^{\alpha\cdot(c-y')}\,dm'. \tag{37}$$

It follows from the above discussion that

$$0<\int [v\cdot(c-y')]^2 e^{\alpha\cdot(c-y')}\,dm'=\sum_{i,j}\psi_{ij}v_i v_j. \tag{38}$$

Hence (ψ_{ij}) is a positive definite matrix and $\psi(\alpha)$ is a "convex" function.

We shall show that $\psi(\alpha)\to\infty$ as $\alpha\to\infty$ in any direction from an arbitrary point α_0 in D'. Choose any unit vector $\alpha_1 (\alpha_1\cdot\alpha_1=1)$, and set $\alpha=\alpha_0+t\alpha_1$, $t\geqq 0$. As t increases, either (A) α reaches the boundary of D', or (B) $t\to\infty$. In case (A), $z'(\alpha)$ becomes unbounded for finite α; hence $\psi(\alpha)$ becomes unbounded. In case (B), since $\widetilde{H}$ is open, there exists an $\varepsilon>0$ such that the point $c-\varepsilon\alpha_1$ is in $\widetilde{H}$. If we consider the hyperplane through that point, $\alpha_1\cdot(c-\varepsilon\alpha_1-y')=\alpha_1\cdot(c-y')-\varepsilon=0$, the half-space $\alpha_1\cdot(c-y')-\varepsilon>0$ must be of positive M'-measure. Therefore

$$m'\{\gamma'\ni\alpha_1\cdot[c-y'(\gamma')]>\varepsilon\}>0.$$

Thus

$$\begin{aligned}\psi(\alpha)&=\int e^{(\alpha_0+t\alpha_1)\cdot(c-y')}\,dm'\geqq e^{\alpha_0\cdot c}\int_{\{\gamma'\ni\alpha_1\cdot[c-y'(\gamma')]>\varepsilon\}} e^{t\alpha_1\cdot[c-y'(\gamma')]-\alpha_0\cdot y'}\,dm'\\ &\geqq e^{\alpha_0\cdot c}e^{t\varepsilon}\int_{\{\gamma'\ni\alpha_1\cdot[c-y'(\gamma')]>\varepsilon\}} e^{-\alpha_0\cdot y'}\,dm'.\end{aligned} \tag{39}$$

We see that $\psi(\alpha)\geqq a_1 e^{t\varepsilon}$ where $a_1>0$ and $\varepsilon>0$. Therefore $\psi(\alpha)\to\infty$ as $t\to\infty$.

It follows that $\psi(\alpha)$ has a minimum at some point α in D'. Since D' is connected, α is unique. The point α is also a minimum for $\log\psi=\alpha\cdot c+\log z'(\alpha)$. Hence, at this point $\frac{\partial}{\partial\alpha_i}\log\psi=c_i-(z')^{-1}\int y'_i e^{-\alpha\cdot y'}\,dm'=0$, $i=1,\ldots,k$; *i.e.*,

$$c=[z'(\alpha)]^{-1}\int y'(\gamma')\,e^{-\alpha\cdot y'(\gamma')}\,dm'=\bar{y}'(\alpha).$$

★ Note that except for the indecomposable case, we can assume that $\widetilde{H}$ is nonempty. This assures us that the corollary is not an empty statement.

2. Applications

A. Initial value problems

The mathematical theory we have developed will be applied to the study of the behavior of a system of the type for which an initial value problem is appropriate. This problem is to determine for all times, t, the "state" of the system when the state at $t=0$ is specified. The state is usually given by a set of numbers or functions. We shall denote the states, whatever they may be, by γ, and the *state space* (the set of all states) by Γ. If we assume the existence and uniqueness of the solution of the initial value problem for all $t \geqq 0$, then we may define a one-parameter family of transformations T_t of Γ into itself depending on the non-negative parameter t: For any $\gamma \in \Gamma$ and any $t \geqq 0$, $T_t \gamma$ is the state at time t which is the solution of the initial value problem corresponding to the state γ at $t=0$. T_t is sometimes called the *solution operator*. It is immediately apparent that for each non-negative t and s, $T_{s+t} = T_s T_t$. In general we are unable actually to solve the initial value problem, but we are interested in obtaining information about the system without knowing the transformations T_t explicitly.

If we can find a completely additive class $\mathscr{A}$ of subsets of Γ, a measure function m, and a complete invariant vector y, such that $(\Gamma, \mathscr{A}, m, T_t, y)$ is a complete space and $m[y^{-1}(I)]$ is finite for every finite interval I in R^k, then we can apply Theorem 1: If $f(\gamma)$ is integrable on (Γ, a, m), then the "infinite time average"

$$f^*(\gamma) = \lim_{L\to\infty} \frac{1}{L} \int_0^L f(T_t \gamma)\, dt \tag{40}$$

exists a.e. and f^* is given by the formula (11). If $f(\gamma)$ represents a physically observable quantity associated with the system in the state γ, we may assume that a measurement of f produces a "finite time average" $\frac{1}{L}\int_0^L f(T_t\gamma)\,dt$. If a large number of repeated measurements of f produce the same value, we may say that the system is in "equilibrium with respect to f", and it is reasonable to identify the measured value with the infinite time average (40). $f^*(\gamma)$ may be called the "equilibrium value of f", and (11) provides a formula for this value which does not require a knowledge of the solution operator T_t.

Physical systems which may be described in terms of an initial value problem are always isolated, *i.e.*, the influence of the physical environment on the system may be neglected. However, in many situations the system of interest* Γ' interacts with the environment to an extent which may not be ignored. In such cases, the solution of the initial value problem no longer describes the behavior of the system. However, it may be possible to consider a larger system, consisting of the system Γ' plus the environment, for which the initial value problem is appropriate, *i.e.*, the combined system is isolated.

If T_t' is the solution operator for the system Γ' when isolated, we assume that we can define a complete space $(\Gamma', \mathscr{A}', m', T_t', y')$. Similarly, we assume

* The symbol Γ' will be used to stand both for the system of interest and for its state space.

that the environment is a physical system Γ'' whose isolated behavior is described by the solution operator T_t'' and that we can define a complete space $(\Gamma'', \mathscr{A}'', m'', T_t'', y'')$. We assume that y' and y'' both have dimension k. Let us suppose, for the moment, that there is no interaction among the systems. Let $\Gamma = \Gamma' \times \Gamma''$, $\mathscr{A} = \mathscr{A}' \times \mathscr{A}''$, $m = m' \times m''$, and let T_t be the natural transformation of Γ induced by T_t' and T_t'', *i.e.*,

$$T_t \gamma = T_t(\gamma', \gamma'') = (T_t' \gamma', T_t'' \gamma''). \tag{41}$$

Then it is easy to show that $(\Gamma, \mathscr{A}, m, T_t)$ is a measure-preserving space and all the invariants y_j', y_j''; $j = 1, \ldots, k$; are invariants of T_t.

Now suppose there is interaction between Γ' and Γ''. Then T_t will no longer be given by (41), nor, in general, will T_t be measure-preserving with respect to $(\Gamma, \mathscr{A}, m,)$ nor will all the y_j', y_j'' be invariants of T_t. However, if the magnitude of the physical interaction is small, we may suppose that some of the above features are preserved. This leads to the definition of "weak interaction" given in the previous secton. We assume that $(\Gamma, \mathscr{A}, m, T_t, y)$ is a complete space where

$$y = \{y_1, \ldots, y_k\} = y' + y''. \tag{42}$$

Some of the requirements of the definition of weak interaction, in particular (42), may be met only in an approximate sense*.

One further physical condition must be met before we can apply our theory: Γ'' must be "large" compared to Γ'. By this we mean that the behavior of the system Γ'' is not appreciably affected by its interaction with Γ'. In particular, the value of $y = y'' + y'$ is not appreciably affected by fluctuations in the value of y'. This condition may be viewed as the physical interpretation of (14), which, in the applications, can be expected to hold only in an asymptotic sense.

Now if $(\Gamma'', \mathscr{A}'', m'', T_t'', y'') \gg (\Gamma', \mathscr{A}', m', T_t', y')$, we may apply Theorem 2: If $f(\gamma')$ is some physically observable quantity associated with the system Γ' in the state γ', the "infinite time average"

$$f^*(\gamma) = \lim_{L \to \infty} \frac{1}{L} \int_0^L f(T_t \gamma)\, dt \tag{43}$$

exists a.e. with respect to m, and (29) provides a simple formula for f^* which does not require a knowledge of T_t: $f^* = \bar{f}(\alpha)$.

But presumably we wish to predict the value of f^* which we may expect to measure in a given situation. For this, we must know the value of α. The significance of the corollary to Theorem 2 is that α can be obtained from a knowledge of $\bar{y}'$, which is the infinite time average of $y'(\gamma')$, and thus is presumably measurable. Knowing $\bar{y}'$, we can then predict that a measurement of f will produce the value $\bar{f}[\alpha(\bar{y}')]$.

Thus the situation is somewhat analogous to the case of the isolated system. There the value of f one would expect to measure was given by f^*, and f^* was uniquely determined by the (invariant) value of y. Here the value of f we expect to measure is given by $\bar{f}$, and $\bar{f}$ is uniquely determined by the value of $\bar{y}'$.

* Thus, *e.g.*, there may be some "interaction parameter", ε, and (42) may be an asymptotic relation valid for $\varepsilon \to 0$.

B. The generalized temperature

Let us suppose that the system of interest, Γ', itself consists of a fixed number, q, of systems in weak interaction. We suppose that q real physical systems are in weak interaction with one another, and the combined system is in weak physical interaction with its environment.

Mathematically, in addition to the hypotheses of Theorem 2, we assume that $(\Gamma_j, \mathscr{A}_j, m_j, T_{jt}, y_{(j)})$ is a complete space for $j=1, \ldots, q$; where $y_{(j)}(\gamma_j) = \{y_{(j)1}, \ldots, y_{(j)k}\}$; $(\Gamma', \mathscr{A}', m') = \Gamma_1 \times \cdots \times \Gamma_q$, $\mathscr{A}_1 \times \cdots \times \mathscr{A}_q$, $m_1 \times \cdots \times m_q)$; and $y' = \sum_{j=1}^{q} y_{(j)}$.

We note that these assumptions imply that the complete spaces $(\Gamma_j, \mathscr{A}_j, m_j, T_{jt}, y_{(j)})$; $j=1, \ldots, q$ are in weak interaction* We also note that $D' = \bigcap_{j=1}^{q} D_j$, where D_j is the set of vectors α in R^k such that $\int e^{-\alpha \cdot y_{(j)}} d m_j < \infty$.

From the conclusions of Theorem 2, we have

$$z'(\alpha) = \int e^{-\alpha \cdot y'} dm' = \prod_{j=1}^{q} z_j(\alpha); \quad z_j(\alpha) = \int e^{-\alpha \cdot y_{(j)}} dm_j. \tag{44}$$

For functions $f(\gamma_1)$ such that $f(\gamma_1) e^{-\alpha \cdot y_{(1)}}$ is integrable on $(\Gamma_1, \mathscr{A}_1, m_1)$ for each α in D_1,

$$\begin{aligned}\bar{f}(\alpha) = [z'(\alpha)]^{-1} \int f(\gamma_1)\, e^{-\alpha \cdot y'(\gamma')} dm' &= (z')^{-1} \int f\, e^{-\alpha \cdot y_{(1)}} dm_1 \cdot \prod_{j=2}^{q} z_j \\ &= (z_1)^{-1} \int f(\gamma_1)\, e^{-\alpha \cdot y_{(1)}} dm_1.\end{aligned} \tag{45}$$

If we omit subscripts "1" associated with the first component of the system of interest, *i.e.* with the space Γ_1, we have

$$z(\alpha) = \int e^{-\alpha \cdot y} dm, \tag{46}$$

$$\bar{f}(\alpha) = [z(\alpha)]^{-1} \int f(\gamma)\, e^{-\alpha \cdot y(\gamma)} dm. \tag{47}$$

Note that although we have singled out the system Γ_1 for special treatment, we could have done the same for any system Γ_j, $j=1, \ldots, q$; with similar results. If we also omit the accent $'$ in (23) and (29), we see that these equations are identical to (46) and (47). Now let us compare the two cases: (I) The system represented by Γ is in weak interaction with its environment. (II) The system represented by Γ is in weak interaction with $q-1$ other systems, and the combined system is in weak interaction with its environment.

We conclude from the above remarks that *$\bar{f}$ is the same in cases* (I) *and* (II) *provided only that α is the same.* Furthermore, suppose we now measure the quantity $f_j(\gamma_j)$ in system Γ_j, for $j=1, \ldots, q$. We would expect in each case to obtain a value $\bar{f}_j(\alpha)$ where α is the same for every j, since we have seen that α is uniquely determined by $\bar{y}'$.

Up to now, we have viewed the parameter α merely as a mathematical convenience. In fact we went to great lengths to show that $\bar{f}$ could be expressed as a function of $\bar{y}'$ rather than as a function of α. However, we see now that

* The definition of "weak interaction" for more than two complete spaces is an obvious extension of our earlier definition.

α may be viewed as a parameter of the system which is the same for systems in weak interaction with one another. From this point of view, α takes on physical significance, and its behavior is very much like a temperature. For this reason, we shall refer to α as the *generalized temperature* of the system.

C. The generalized "Second Law"

Let $\beta=(\beta_1, \dots, \beta_r)$ be a vector in R^r, and let us assume that y' and T_t' depend parametrically on β. In addition, we assume that, for each fixed value of β in some region I_β of R^r, all the hypotheses of Theorem 2 hold; hence all the conclusions also hold. Then our main results are summarized in the formulas (we omit the accent)

(48) $$z = z(\alpha;\beta) = \int e^{-\alpha\cdot y(\gamma;\beta)}\,dm,$$

(49) $$\bar{f} = \bar{f}(\alpha;\beta) = z^{-1}\int f(\gamma)\,e^{-\alpha\cdot y(\gamma;\beta)}\,dm.$$

If we assume that we can differentiate formally, we have

(50) $$d(\log z) = -z^{-1}\int e^{-\alpha\cdot y}\left\{\sum_i y_i\,d\alpha_i + \sum_i\left[\sum_k \alpha_k\,\frac{\partial y_k}{\partial\beta_i}\right]d\beta_i\right\}dm,$$

(51) $$d(\log z) = -\sum_i \bar{y}_i\,d\alpha_i - \sum_i\left[\sum_k \alpha_k\overline{\left(\frac{\partial y_k}{\partial\beta_i}\right)}\right]d\beta_i.$$

Let us define

(52) $$q = q(\alpha,\beta,\gamma) = z^{-1}e^{-\alpha\cdot y}; \qquad H = H(\alpha,\beta) = \int q\log q\,dm.$$

Then $\int q\,dm = 1$ and $\log q + \log z + \alpha\cdot y = 0$. If we multiply this equation by q and integrate with respect to m over Γ we obtain

(53) $$H + \log z + \alpha\cdot\bar{y} = 0.$$

Differentiating (53) we have $dH + d(\log z) + \sum_i \alpha_i\,d(\bar{y}_i) + \sum_i \bar{y}_i\,d\alpha_i = 0$. Combining this with (51) yields

(54) $$-dH = \sum_i \alpha_i\,d(\bar{y}_i) - \sum_i\left[\sum_k \alpha_k\overline{\left(\frac{\partial y_k}{\partial\beta_i}\right)}\right]d\beta_i.$$

This equation will be called the *generalized second law* because of its similarity to the second law of Thermodynamics. Indeed we shall compare it directly with that law shortly. For later reference we shall also need the following equations which are obtained from (48) by differentiation:

(55) $$\bar{y}_k = -\frac{\partial}{\partial\alpha_k}\log z,$$

(56) $$\sum_k \alpha_k\overline{\left(\frac{\partial y_k}{\partial\beta_i}\right)} = -\frac{\partial}{\partial\beta_i}\log z.$$

D. Classical mechanics

A mechanical system with s degrees of freedom may be described in terms of generalized coordinates $q_1, \dots, q_s$ and conjugate momenta $p_1, \dots, p_s$ which are solutions of HAMILTON's equations

(57) $$\frac{dq_i}{dt} = \frac{\partial E}{\partial p_i}, \qquad \frac{dp_i}{dt} = -\frac{dE}{\partial q_i}; \qquad i = 1, \dots, s.$$

Here $E=E(q_1, \ldots, q_s, p_1, \ldots, p_s, t)$ is the total energy or "Hamiltonian" of the system. We shall denote the $2s$ coordinates and momenta by a single vector $\varrho=\{\varrho_1, \ldots, \varrho_{2s}\}=\{q_1, \ldots, q_s, p_1, \ldots, p_s\}$ which may be thought of as a point in $2s$-dimensional euclidean space or "phase space", P. We shall assume that the mechanical system is "conservative", *i.e.*, E is independent of t, and that E depends on r parameters $\beta_1, \ldots, \beta_r$. Thus $E=E(\varrho;\beta)$ where $\beta=\{\beta_1, \ldots, \beta_r\}$, and β varies in some region I_β of R^r. We assume that as a function of (ϱ, β) defined in $P \times I_\beta$, E has the necessary properties such that the initial value problem defined by (57) and

$$\varrho=\varrho^0 \quad \text{at} \quad t=0 \tag{58}$$

has a unique solution $\varrho(t;\varrho^0,\beta)$ which has continuous first derivatives with respect to all its arguments in $R^1 \times P \times I_\beta$. (Here R^1 is the set $-\infty<t<\infty$.) If we define the solution operator T_t by

$$T_t\varrho^0=\varrho(t;\varrho^0,\beta); \qquad -\infty<t<\infty, \quad \varrho^0 \in P, \quad \beta \in I_\beta, \tag{59}$$

then for each fixed t and β, T_t is a continuously differentiable transformation of phase space into itself.

Now for any real s, the two functions $\varrho(t;T_s\varrho^0,\beta)$ and $\varrho(t+s;\varrho^0,\beta)$ are both equal to $T_s\varrho^0$ at $T=0$. It follows from the uniqueness assertion that the two functions are equal and hence that

$$T_{t+s}=T_t T_s \qquad \text{for all real } t \text{ and } s. \tag{60}$$

In particular $T_{-t}T_t=T_0=I$, where I is the identity transformation. Thus for every t, T_t has an inverse, and $T_t^{-1}=T_{-t}$.

It follows easily from (57) that

$$\begin{aligned} \frac{d}{dt}E[q_i(t),p_i(t);\beta] &= \sum_i \left[\frac{\partial E}{\partial q_i}\frac{dq_i}{dt}+\frac{\partial E}{\partial p_i}\frac{dp_i}{dt}\right] \\ &= \sum_i \left[\frac{\partial E}{\partial q_i}\frac{\partial E}{\partial p_i}-\frac{\partial E}{\partial p_i}\frac{\partial E}{\partial q_i}\right]=0. \end{aligned} \tag{61}$$

Thus $E[\varrho(t);\beta]=E[\varrho(0);\beta]$, or

$$E[T_t\varrho^0;\beta]=E[\varrho^0;\beta] \qquad \text{for all } t \in R^1 \text{ and } \varrho^0 \in P. \tag{62}$$

The following Lemma contains assertions which are needed in order to apply our theory to mechanical systems. The proof appears in Appendix 1.

Lemma 6. *Let μ denote Lebesgue measure, and $\mathscr{L}$ denote the class of Lebesgue measurable sets in phase space P. Let $\mathscr{M}$ denote the class of Lebesgue measurable sets in R^1. Then*

(I) *The transformations T_t are measure-preserving with respect to $(P, \mathscr{L}, \mu)$.*

(II) *$T_t\varrho$ is a measurable function of t and ϱ,* i.e., *$T_t\varrho$ is a measurable transformation from $(R^1 \times P, \mathscr{M} \times \mathscr{L})$ into $(P, \mathscr{L})$.*

(III) *E is a measurable function with respect to $(P, \mathscr{L}, \mu)$.*

From (III) and (62) it follows that E *is an invariant function* for each fixed β, and from (I), (II) and (60) that $(P, \mathscr{L}, \mu, T_t)$ *is a measure-preserving space.*

We now restrict our considerations to mechanical systems for which the function E is, by itself, a complete set of invariants*. Experimental evidence indicates that this assertion leads to consistent results for a surprisingly large class of systems. Thus $(P, \mathscr{L}, \mu, T_t, E)$ *is a complete space*. For many Hamiltonians of practical interest, it can be verified that for every finite interval, I, in R^1,

$$\mu\{E^{-1}(I)\} = \mu\{\varrho \ni E(\varrho) \in I\} < \infty. \tag{63}$$

We restrict consideration to mechanical systems satisfying this condition, also.

For isolated mechanical systems satisfying the above conditions, we may now apply Theorem 1. If $f(\varrho)$ is a "phase function", *i.e.*, if f is integrable on $(P, \mathscr{L}, \mu)$, then the infinite time average $f^*(\varrho)$ exists a.e. If the system is in equilibrium with respect to f, we may predict that a measurement of f will produce the value f^*, and we have the following formula:

$$f^*[E^{-1}(x)] = \lim_{j\to\infty} \frac{\int\limits_{\{\varrho \ni c^j < E(\varrho;\beta) - x \leqq d^j\}} f(\varrho)\, d\mu}{\{\varrho \ni c^j < E(\varrho;\beta) - x \leqq d^j\}}. \tag{64}$$

$f^*[E^{-1}(x)]$ denotes the value of $f^*(\varrho)$ common to all ϱ for which $E(\varrho) = x$; c^j, d^j are sequences of numbers, which depend on x, for which $c^j < 0 \leqq d^j$, $\lim\limits_{j\to\infty} (d^j - c^j) = 0$. Formula (64) is sometimes referred to as the *microcanonical distribution.*

If our mechanical system interacts physically with its environment, we assume that we may treat the system and environment together as a mechanical system whose states are points in a phase space $P = P' \times P''$, where P' is the phase space of the system of interest and P'' is the phase space of the environment. It is certainly true that the transformations T_t of P determined by HAMILTON's equations and the total energy E are measure-preserving with respect to $(P, \mathscr{L}, \mu)$, where μ is Lebesgue measure in P, and that E is an invariant of T_t. We again restrict consideration to systems for which E is a complete set of invariants. If the magnitude of the physical interaction is small, then E will be approximately equal to the sum of the energies of the component systems. Thus the condition

$$E = E' + E'' \tag{65}$$

will be satisfied in an asymptotic sense. It is easy to show that $(P, \mathscr{L}, \mu) = (P' \times P'', \mathscr{L}' \times \mathscr{L}'', \mu' \times \mu'')$. Thus we may assume that the complete spaces $(P', \mathscr{L}', \mu', T_t, E')$, $(P'', \mathscr{L}'', \mu'', T_t'', E'')$ are in weak interaction, according to our definition of that term. For most Hamiltonians E' of practical interest, the set D' of all real numbers α such that

$$z'(\alpha) = \int\limits_{P'} e^{-\alpha E'(\varrho')}\, d\mu' < \infty, \tag{66}$$

consists of the set $\{0 < \alpha < \infty\}$, which is an open connected set. Thus, in order to apply the conclusions of Theorem 2 and its corollary, we need only assume that the environment is "large" in the sense defined.

---

* For a discussion of certain mechanical systems where E alone is not a complete set of invariants, see [2].

Let us now examine the consequences of Theorem 2. To conform with general usage we shall denote α by ϑ, and shall omit the accent. For ϑ in $D=\{0<\vartheta<\infty\}$,

$$z(\vartheta,\beta)=\int_P e^{-\vartheta E(\varrho,\beta)}\,d\mu. \tag{67}$$

If $f(\varrho)\,e^{-\vartheta E(\varrho,\beta)}$ is integrable on $(P,\mathscr{L},\mu)$ for each ϑ in D, then

$$\bar{f}(\vartheta,\beta)=z^{-1}\int_P f(\varrho)\,e^{-\vartheta E(\varrho,\beta)}\,d\mu. \tag{68}$$

Formulas (67) and (68) are sometimes referred to as the *canonical distribution*.

As in Section 2C, we define

$$q=z^{-1}\,e^{-\vartheta E}, \tag{69}$$

and

$$H=\int q\log q\,d\mu. \tag{70}$$

Then from (53)

$$H+\log z+\vartheta\,\bar{E}=0. \tag{71}$$

From (54)

$$-dH=\vartheta\,d(\bar{E})-\vartheta\sum_i\overline{\left(\frac{\partial E}{\partial\beta_i}\right)}\,d\beta_i. \tag{72}$$

Systems of the type we have been discussing are studied in thermodynamics. The main result of this study is the "second law of thermodynamics" which may be written as

$$T\,dS=dU+\sum_i p_i\,d\beta_i, \tag{73}$$

where p_i denotes the "generalized force" corresponding to the "external variable" β_i, T denotes absolute temperature, U internal energy, and S entropy. Comparing (72) and (73), we see that these equations are identical if we set

$$\frac{1}{\vartheta}=kT;\quad -kH=S;\quad \bar{E}=U;\quad -\overline{\left(\frac{\partial E}{\partial\beta_i}\right)}=p_i. \tag{74}$$

The constant k can have any real value. By convention it is fixed at a certain value, and this determines a convenient temperature scale. In terms of the identities (74), z, which is called the *partition function*, may be written

$$z=\int_P e^{-\frac{1}{kT}E(q_i,p_i;\beta)}\,dq_i\,dp_i. \tag{75}$$

From (71) we have

$$-kT\log z=U-TS. \tag{76}$$

In thermodynamics the quantity $U-TS$ is called the "free energy" and is denoted by A. Thus

$$A=-kT\log z=-\frac{1}{\vartheta}\log z. \tag{77}$$

From (55) we obtain

$$U=kT^2\frac{\partial}{\partial T}\log z=-T^2\frac{\partial}{\partial T}\left(\frac{A}{T}\right), \tag{78}$$

and from (56)

$$p_i = kT\frac{\partial}{\partial \beta_i}\log z = -\frac{\partial A}{\partial \beta_i}. \tag{79}$$

Equations (76), (77), (78), and (79) express the thermodynamic variables p_i, U, A, and S as functions of T and β which involve only the partition function and its logarithmic derivatives.

E. The grand canonical distribution

We now consider a mechanical system consisting of δ different kinds of particles. If n_i denotes the number of particles of the i^{th} kind, and s_i the number of degrees of freedom of the i^{th} kind of particle, then the states of the system can be represented in a phase space with points $\{q_{ijk}, p_{ijk}\}$ where $i=1, \ldots, \delta$ labels the kind of particle; $j=1, \ldots, n_i$ numbers the particles of that kind; and $k=1, \ldots, s_i$ determines the coordinates corresponding to each degree of freedom. The vector $n=\{n_1, \ldots, n_\delta\}$ determines the "composition" of the system, *i.e.*, the number of particles of each kind.

For a given composition, n, we have a phase space P_n of points $\varrho_n=(q_{ijk}, p_{ijk})$. We assume that as a function of ϱ_n, β defined in $P_n \times I_\beta$, the Hamiltonian $E_n(\varrho_n, \beta)$ has the necessary properties such that the initial value problem has a unique continuously differentiable solution. If μ_n denotes Lebesgue measure, $\mathscr{L}_n$ denotes the class of Lebesgue measurable sets in P_n, and if S_{nt} denotes the solution operator for the system, then we know from the discussion of the preceding section that $(P_n, \mathscr{L}_n, \mu_n, S_{nt})$ is a measure-preserving space and E_n is an invariant function. As before we shall restrict consideration to systems for which E_n is, by itself, a complete set of invariants. Thus $(P_n, L_n, \mu_n, S_{nt}, E_n)$ is a complete space.

As in the preceding section, we shall assume that our physical system interacts with its environment, but the interaction will include not only the energy, but the particles themselves. Thus we shall suppose that particles can be exchanged between the system and its environment. In such a situation it is no longer possible to represent the state of the system by a point in a single phase space P_n, but rather as a point in $\bigcup\limits_n P_n$. In this connection, the following Theorem will be used. The proof appears in Appendix 2.

Theorem 3. *Let $\mathscr{N}$ be the space of vectors $n=\{n_1, \ldots, n_\delta\}$ where n_i is a non-negative integer for $i=1, \ldots, \delta$. For each $n \in \mathscr{N}$, let $(P_n, \mathscr{L}_n, \mu_n, S_{nt}, g_{(n)})$ be a complete space, where $g_{(n)}=g_{(n)1}, \ldots, g_{(n)k}=g_{(n)}(\varrho_n)$, and ϱ_n denotes a point in P_n. Let*

$$\Gamma = \bigcup_{n\in\mathscr{N}} P_n. \tag{80}$$

Let $\mathscr{A}$ be the class of subsets A of Γ such that

$$A \cap P_n \in \mathscr{L}_n \quad \textit{for every } n \in \mathscr{N}. \tag{81}$$

For $A \in \mathscr{A}$, let

$$m(A) = \sum_{n\in\mathscr{N}} \mu_n(A \cap P_n). \tag{82}$$

For each $\gamma = \varrho_n$ in Γ, let

$$T_t\gamma = S_{nt}\varrho_n. \tag{83}$$

Let $n(\gamma) = [n_1, \ldots, n_\delta]$ *be the vector function defined by*

$$n(\gamma) = n(\varrho_n) = n, \tag{84}$$

and let

$$g(\gamma) = g(\varrho_n) = g_{(n)}(\varrho_n). \tag{85}$$

Let $y(\gamma)$ *be the* $(\delta + k)$*-component vector function defined by*

$$y(\gamma) = \{n(\gamma), g(\gamma)\}. \tag{86}$$

Then $(\Gamma, \mathscr{A}, m, T_t, y)$ *is a complete space.* $f(\gamma) = f_n(\varrho_n)$ *is a measurable function with respect to* $(\Gamma, \mathscr{A})$ *if and only if* $f_n(\varrho_n)$ *is a measurable function with respect to* $(P_n, \mathscr{L}_n)$ *for every* $n \in \mathscr{N}$. *If* $f(\gamma)$ *is integrable on* $(\Gamma, \mathscr{A}, m)$ *then* $f_n(\varrho_n) = f(\varrho_n)$ *is integrable on* $(P_n, \mathscr{L}_n, \mu_n)$ *for each* $n \in \mathscr{N}$,

$$\int_\Gamma f(\gamma)\, dm = \sum_{n \in \mathscr{N}} \int_{P_n} f_n(\varrho_n)\, d\mu_n, \tag{87}$$

and the series converges absolutely. Conversely if $f_n(\varrho_n)$ *is non-negative and integrable on* $(P_n, \mathscr{L}_n, \mu_n)$ *for each* $n \in \mathscr{N}$, *and if the above series converges, then* $f(\gamma) = f_n(\varrho_n)$ *is integrable on* $(\Gamma, \mathscr{A}, m)$ *and* (87) *holds.*

We now apply this theorem setting $g_{(n)}(\varrho_n) = E_n(\varrho_n)$ (thus $k=1$). In order to apply Theorem 2 we set $(\Gamma', \mathscr{A}', m', T_t', y') = (\Gamma, \mathscr{A}, m, T_t, y)$, and assume that the environment can be represented by a similar space $(\Gamma'', \mathscr{A}'', m'', T_t'', y'')$ where $(\Gamma'', \mathscr{A}'', m'', T_t'', y'') \gg (\Gamma', \mathscr{A}', m', T_t', y')$. The vector α will be denoted by $\{\lambda, \vartheta\}$ where $\lambda = \{\lambda_1, \ldots, \lambda_\delta\}$ and $\alpha \cdot y = \lambda \cdot n + \vartheta E_n$. In stating the results, we omit the accent: For $\{\lambda, \vartheta\}$ in some subset D of $R^{\delta+1}$ (which is assumed to be a non-empty, open, connected set)

$$z(\lambda, \vartheta, \beta) = \int e^{-\alpha \cdot y}\, dm = \sum_{n \in \mathscr{N}} e^{-\lambda \cdot n} \int_{P_n} e^{-\vartheta E_n(\varrho_n, \beta)}\, d\mu_n. \tag{88}$$

If $f(\gamma)\, e^{-\alpha \cdot y(\gamma)} = f_n(\varrho_n)\, e^{-\lambda \cdot n - \vartheta E_n(\varrho_n)}$ is integrable on $(\Gamma, \mathscr{A}, m)$ for every $\alpha = \{\lambda, \vartheta\}$ in D, then

$$\bar{f}(\lambda, \vartheta, \beta) = z^{-1} \int f\, e^{-\alpha \cdot y}\, dm = z^{-1} \sum_{n \in \mathscr{N}} e^{-\lambda \cdot n} \int_{P_n} f_n(\varrho_n)\, e^{-\vartheta E_n(\varrho_n)}\, d\mu_n. \tag{89}$$

Formulas (88) and (89) are sometimes referred to as the *grand canonical distribution.*

As in Section 2C we define

$$q = z^{-1} e^{-\lambda \cdot n - \vartheta E_n}, \tag{90}$$

$$H = \int q \log q\, dm. \tag{91}$$

Then from (53)

$$H + \log z + \lambda \cdot \bar{n} + \vartheta \bar{E}_n = 0. \tag{92}$$

From (54)

$$-dH = \vartheta\, d\bar{E}_n - \vartheta \sum_i \overline{\left(\frac{\partial E_n}{\partial \beta_i}\right)}\, d\beta_i + \sum_i \lambda_i\, d\bar{n}_i. \tag{93}$$

Systems of the type which we have been describing are also studied in thermodynamics. For this type of system the second law may be written

$$T\, dS = dU + \sum_i p_i\, d\beta_i - \sum_i \mu_i\, d\nu_i. \tag{94}$$

μ_i is the "partial free energy" or "partial thermal potential" associated with the i^{th} kind of particle, and ν_i is the average number of particles of the i^{th} kind. Comparing (93) and (94), we see that these equations are identical if we set

$$\frac{1}{\vartheta} = kT; \quad -kH = S; \quad \bar{E}_n = U; \quad -\overline{\left(\frac{\partial E_n}{\partial \beta_i}\right)} = p_i; \quad \frac{\lambda_i}{\vartheta} = -\mu_i; \quad \bar{n}_i = \nu_i. \tag{95}$$

In terms of these identities, z, which is called the *grand partition function* may be written

$$z = \sum_{n_1, \ldots, n_\delta = 0}^{\infty} \exp\left\{\sum_i \frac{\mu_i n_i}{kT}\right\} \int_{P_n} \exp\left\{-\frac{1}{kT} E_n(q_{ijk}, p_{ijk}; \beta)\right\} dq_{ijk}\, dp_{ijk}. \tag{96}$$

From (92) we have

$$-kT \log z = U - TS - \sum_i \mu_i \nu_i. \tag{97}$$

From (55) we obtain

$$U = kT^2 \frac{\partial}{\partial T} \log z \tag{98}$$

and

$$\nu_i = kT \frac{\partial}{\partial \mu_i} \log z, \tag{99}$$

and from (56)

$$p_i = kT \frac{\partial}{\partial \beta_i} \log z. \tag{100}$$

Equations (97), (98), (99), and (100) express the thermodynamic variables p_i, ν_i, U, S as functions of T, μ_i, and β which involve only the grand partition function and its logarithmic derivatives.

Appendix 1: Proof of Lemma 6

(I) *The transformations T_t are measure-preserving with respect to $(P, \mathscr{L}, \mu)$.*

To prove (I) it must be shown that for every $L \in \mathscr{L}$ and for every t, $T_t^{-1} L \in \mathscr{L}$ and $\mu(T_t^{-1} L) = \mu(L)$. The proof depends on a famous theorem of of LIOUVILLE*.

A. For every fixed t, the Jacobian of the transformation T_t is everywhere equal to 1.

B. We shall use the following definition of Lebesgue measure**: Let C be the class consisting of the empty set φ and the open intervals $\{a < \varrho < b\} = \{\varrho \ni a_i < \varrho_i < b_i;\ i = 1, \ldots, 2s\}$ in P. For any set $G \in C$, let $\tau(G)$ denote the $2s$-dimensional volume of G. For any subset A of P, set

$$\mu^*(A) = \inf\left\{\sum_{n=1}^{\infty} \tau(G_n) \ni G_n \in C;\ \bigcup_{n=1}^{\infty} G_n > A\right\}. \tag{A 1.1}$$

μ^* is a set function defined on the class of all subsets of P and is called *Lebesgue outer measure*. By definition, $\mathscr{L}$ is the class of all sets L such that

$$\mu^*(A) = \mu^*(A \cap L) + \mu^*(A - L) \qquad \text{for every } A < P; \tag{A 1.2}$$

and μ is the restriction of μ^* to the sets of $\mathscr{L}$.

* See [*1*].

** See [*6*], p. 93.

C. Let $\mathcal{D}$ be the class of all continuously differentiable transformations of P with Jacobian everywhere equal to 1. Let C_1 be the class of all subsets E_1 of P such that $E_1 = TE$ for some $T \in \mathcal{D}$ and some $E \in C$. For every set $E_1 \in C_1$ let $\tau(E_1)$ be the $2s$-dimensional volume of E_1. (Then $\tau(E_1) = \tau(E)$.) For any subset A of P, set

$$\mu_1^*(A) = \inf\left\{\sum_{n=1}^{\infty} \tau(E_n) \ni E_n \in C_1;\ \bigcup_{n=1}^{\infty} E_n > A\right\}. \tag{A 1.3}$$

We shall show that $\mu_1^*(A) \equiv \mu^*(A)$. In so doing we shall make use of the following theorem*:

D. Every open set in P is the union of a countable class of disjoint, half-open intervals.

E. For every set E_1 in C_1, and every $\varepsilon > 0$, there exists a sequence of sets G^k such that $G^k \in C$, $U_k G^k > E_1$, and $\sum_k \tau(G^k) \leqq \tau(E_1) + \varepsilon$:

If F is any half-open interval $(a, b]$, let $\tau(F)$ denote the $2s$-dimensional volume of F, *i.e.*, $\tau(F) = \prod_{i=1}^{2s} (b_i - a_i)$. Given $\delta > 0$, there exists an open interval $G = (a, c)$ such that $F < G$ and $\tau(G) = \prod_{i=1}^{2s} (c_i - a_i) \leqq \tau(F) + \delta$. By D, there exists a sequence of disjoint, half-open intervals F^k whose union is E_1. For each F^k choose an open interval G^k such that $F^k < G^k$ and $\tau(G^k) \leqq \tau(F^k) + \frac{\varepsilon}{2k}$. Then $\bigcup_k G^k > \bigcup_k F^k = E_1$, and $\sum_k \tau(G^k) \leqq \varepsilon + \sum_k \tau(F^k) = \varepsilon + \tau(E_1)$.

F. $\mu_1^*(A) = \mu^*(A)$ for every $A < P$:

Since $C < C_1$, it follows that for every $A < P$

$$\mu_1^*(A) \leqq \mu^*(A). \tag{A 1.4}$$

Given any sequence of sets $E_n \in C_1$ such that $\bigcup_n E_n > A$, and given any $\varepsilon > 0$, we may, by E, choose sets $G_n^k \in C$ such that for each n $\bigcup_{k=1}^{\infty} G_n^k > E_n$ and $\sum_k \tau(G_n^k) \leqq \tau(E_n) + \frac{\varepsilon}{2n}$. Then $\bigcup_{n,k} G_n^k > A$ and $\sum_{n,k} \tau(G_n^k) \leqq \sum_n \left[\tau(E_n) + \frac{\varepsilon}{2n}\right] \leqq \varepsilon + \sum_n \tau(E_n)$. If we renumber the G_n^k with a single index, we have $\bigcup_n G_n > A$ and $\sum_n \tau(G_n) \leqq \varepsilon + \sum_n \tau(E_n)$. Since ε was arbitrary, it follows that

$$\mu_1^*(A) \geqq \mu^*(A). \tag{A 1.5}$$

G. For every $L \in \mathcal{L}$ and every t, $T_t^{-1} L \in \mathcal{L}$ and $\mu(T_t^{-1} L) = \mu(L)$:

For any $A < P$

$$\begin{aligned}\mu^*(A) = \mu_1^*(A) &= \inf\left\{\sum_{n=1}^{\infty} \tau(E_n) \ni E_n \in C_1;\ \bigcup_{n=1}^{\infty} E_n > A\right\} \\ &= \inf\left\{\sum_{n=1}^{\infty} \tau(T_t^{-1} E_n) \ni T_t^{-1} E_n \in C_1;\ \bigcup_{n=1}^{\infty} T_t^{-1} E_n > T_t^{-1} A\right\} \\ &= \mu_1^*(T_t^{-1} A) = \mu^*(T_t^{-1} A).\end{aligned} \tag{A 1.6}$$

* See [6], p. 126.

Let L be any set in $\mathscr{L}$. For any set $B \subset P$ let $A = T_t B$. Then $T_t^{-1} A = B$ and

$$\begin{aligned} \mu^*(B) &= \mu^*(T_t^{-1} A) = \mu^*(A) = \mu^*(A \cap L) + \mu^*(A - L) \\ &= \mu^*[T_t^{-1}(A \cap L)] + \mu^*[T_t^{-1}(A - L)] \\ &= \mu^*[T_t^{-1} A \cap T_t^{-1} L] + \mu^*[T_t^{-1} A - T_t^{-1} L] \\ &= \mu^*(B \cap T_t^{-1} L) + \mu^*(B - T_t^{-1} L). \end{aligned} \tag{A 1.7}$$

Hence $T_t^{-1} L \in \mathscr{L}$ and

$$\mu(T_t^{-1} L) = \mu^*(T_t^{-1} L) = \mu^*(L) = \mu(L). \tag{A 1.8}$$

II. *$T_t \varrho$ is a measurable function of t and ϱ, i.e., $T_t \varrho$ is a measurable transformation from $(R^1 \times P, \mathscr{M} \times \mathscr{L})$ into $(P, \mathscr{L})$:*

Given a measure space $(\Gamma, \mathscr{A}, m)$, let Z be the class of all subsets of sets of measure zero. Let $\bar{\mathscr{A}}$ be the class of sets of the form $z \cup A$, where $z \in Z$ and $A \in \mathscr{A}$. Then $\bar{\mathscr{A}}$ is a completely additive class of sets containing $\mathscr{A}$ and the extension $\bar{m}$ of m to $\bar{\mathscr{A}}$ defined by

$$\bar{m}(z \cup A) = m(A) \tag{A 1.9}$$

is a measure. $\bar{m}$ is called the *completion* of m.

Let ν denote Lebesgue measure on $(R^1, \mathscr{M})$. It can be shown that the Borel sets in R^k are a subclass of the Lebesgue measurable sets, and the restriction of Lebesgue measure to the Borel sets is a measure whose completion is Lebesgue measure. Thus if $\hat{\mathscr{M}}$ denotes the Borel sets in R^1 and $\hat{\nu}$ denotes the restriction of ν to $\hat{\mathscr{M}}$, then $\bar{\hat{\mathscr{M}}} = \mathscr{M}$ and $\bar{\hat{\nu}} = \nu$. We shall also denote the Borel sets in P by $\hat{\mathscr{L}}$ and the restriction of μ to $\hat{\mathscr{L}}$ by $\hat{\mu}$. As before, $\bar{\hat{\mathscr{L}}} = \mathscr{L}$ and $\bar{\hat{\mu}} = \mu$.

It can easily be shown that if f is a measurable transformation from $(X, \mathscr{A}, \sigma)$ into $(Y, \mathscr{B}, \tau)$ such that $\sigma[f^{-1}(B)] = 0$ whenever $\tau(B) = 0$, then f is a measurable transformation from $(X, \bar{\mathscr{A}}, \bar{\sigma})$ into $(Y, \bar{\mathscr{B}}, \bar{\tau})$. Now $f(t, \varrho) = T_t \varrho$ is a continuous transformation from $R^1 \times P$ into P. From this it can be shown that f is measurable from $(R_1 \times P, \hat{\mathscr{M}} \times \hat{\mathscr{L}}, \hat{\nu} \times \hat{\mu})$ into $(P, \hat{\mathscr{L}}, \hat{\mu})$. Thus it will follow that f is measurable from $(R_1 \times P, \mathscr{M} \times \mathscr{L}, \nu \times \mu)$ into $(P, \mathscr{L}, \mu)$ if we can show that

$$\hat{\nu} \times \hat{\mu}[f^{-1}(L)] = 0 \qquad \text{whenever } \hat{\mu}[L] = 0. \tag{A 1.10}$$

Now if G is any subset of $R_1 \times P$, the set $G_t = \{\varrho \ni (t, \varrho) \in G\}$ is a subset of P called the *section of G determined by t*. We shall show that every section $[f^{-1}(L)]_t$ has $\hat{\mu}$-measure zero. From this it follows* that $\hat{\nu} \times \hat{\mu}[f^{-1}(L)] = 0$. Let $f_t(\varrho) = f(t, \varrho) = T_t \varrho$. We have seen that for each fixed t, f_t is a continuously differentiable transformation of P into itself with Jacobian 1. Now

$$[f^{-1}(L)]_t = \{\varrho \ni f(t, \varrho) \in L\} = \{\varrho \ni f_t(\varrho) \in L\} = f_t^{-1}(L). \tag{A 1.11}$$

If L is any Borel set of measure zero and ϵ is any positive number, there exists an open set $\mathcal{O}$ of volume ϵ containing L. Then $f_t^{-1}(\mathcal{O})$ is also of volume ϵ and contains $f_t^{-1}(L)$. Since ϵ was arbitrary, it follows that $\hat{\mu}\{f_t^{-1}(L)\} = \hat{\mu}\{[f^{-1}(L)]_t\} = 0$, and the proof is complete.

---

* See, *e.g.*, [*3*], Sec. 36, Theorem A.

III. *E is a measurable function with respect to* $(P, \mathscr{L}, \mu)$.

This is a simple consequence of the fact that E is a continuous function on P (see [6], p. 146).

Appendix 2. Proof of Theorem 3 *

1. $\mathscr{A}$ is a completely additive class of subsets of Γ.

$\varphi \in \mathscr{A}$. Let $A \in \mathscr{A}$; then $(\Gamma - A) \cap P_n = P_n - A \cap P_n \in \mathscr{L}_n$ for every $n \in \mathscr{N}$; therefore, $(\Gamma - A) \in \mathscr{A}$. Let $A_1, A_2, \ldots$ be in $\mathscr{A}$ and set $A = \bigcup_{k=1}^{\infty} A_k$. Then $A \cap P_n = \left(\bigcup_k A_k \cap P_n\right) \in \mathscr{L}_n$ for every $n \in \mathscr{N}$; therefore, $A \in \mathscr{A}$.

2. m is a σ-finite measure function defined on $\mathscr{A}$:

Let $A_1, A_2, \ldots$ be a sequence of disjoint sets in $\mathscr{A}$. Then

$$m\left(\bigcup_k A_k\right) = \sum_{n \in \mathscr{N}} \mu_n \left(\bigcup_k A_k \cap P_n\right) = \sum_n \sum_k \mu_n (A_k \cap P_n) = \sum_k \sum_n \mu_n (A_k \cap P_n) = \sum_k m(A_k).$$

Let $P_n = \bigcup_{k=1}^{\infty} L_{nk}$; $\mu_n(L_{nk}) < \infty$. Then $\Gamma = \bigcup_n \bigcup_k L_{nk}$, and $m(L_{nk}) = \mu_n(L_{nk}) < \infty$; hence m is σ-finite.

3. $f(\gamma) = f_n(\varrho_n)$ is a measurable function with respect to $(\Gamma, \mathscr{A})$ if and only if $f_n(\varrho_n)$ is a measurable function with respect to $(P_n, \mathscr{L}_n)$ for every $n \in \mathscr{N}$:

For any Borel set $B \subset R^1$, $f^{-1}(B) \cap P_n = \{\gamma \ni f(\gamma) \in B\} \cap P_n = \{\varrho_n \ni f_n(\varrho_n) \in B\} = f_n^{-1}(B)$. Hence $f^{-1}(B) \in \mathscr{A}$ if and only if $f_n^{-1}(B) \in \mathscr{L}_n$ for every $n \in \mathscr{N}$.

4. Let $f(\gamma) = \sum_{k=1}^{p} a_k C_{A_k}(\gamma)$ be a simple function**. Then $f_n(\varrho_n) = f(\varrho_n)$ is simple for each $n \in \mathscr{N}$. If $f(\gamma)$ is integrable on $(\Gamma, \mathscr{A}, m)$, then $f_n(\varrho_n)$ is integrable on $(P_n, \mathscr{L}_n, \mu_n)$ for each $n \in \mathscr{N}$,

$$\int_\Gamma f(\gamma)\, dm = \sum_{n \in \mathscr{N}} \int_{P_n} f_n(\varrho_n)\, d\mu_n, \tag{A 2.1}$$

and the series converges. Conversely, if $f_n(\varrho_n)$ is integrable on $(P_n, \mathscr{L}_n, \mu_n)$ for each $n \in \mathscr{N}$ and the series converges, then $f(\gamma)$ is integrable on $(\Gamma, \mathscr{A}, m)$:

$$\begin{aligned} \int_\Gamma f(\gamma)\, dm &= \sum_k a_k\, m(A_k) = \sum_k a_k \sum_n \mu_n (A_k \cap P_n) = \sum_k a_k \sum_n \int_{P_n} C_{A_k}(\varrho_n)\, d\mu_n \\ &= \sum_n \int_{P_n} f_n(\varrho_n)\, d\mu_n. \end{aligned} \tag{A 2.2}$$

Assertion 4 now follows from (A 2.2) and the definition of integrability for simple functions.

5. Let $f(\gamma)$ be non-negative and measurable. Then $f_n(\varrho_n) = f(\varrho_n)$ is non-negative and measurable for each $n \in \mathscr{N}$. If $f(\gamma)$ is integrable, then $f_n(\varrho_n)$ is integrable for each $n \in \mathscr{N}$, (A 2.1) is valid, and the series converges:

If f is integrable, then, by definition, there exists a non-decreasing sequence $f_1, f_2, \ldots$ of non-negative simple functions, each integrable on $(\Gamma, \mathscr{A}, m)$ such

* The definition of the integral and other measure-theoretic concepts used here are based on [6].

** $C_{A_k}(\gamma)$ denotes the characteristic function of the set A_k. For the definition of a simple function see [6].

that, for each $\gamma \in \Gamma$, $\lim_{k\to\infty} f_k(\gamma) = f(\gamma)$, and such that $\lim_{k\to\infty} \int_\Gamma f_k \, dm < \infty$. By definition

$$\int_\Gamma f(\gamma)\, dm = \lim_{k\to\infty} \int_\Gamma f_k(\gamma)\, dm. \tag{A 2.3}$$

Set $C_{kn} = \int_{P_n} f_k(\varrho_n)\, d\mu_n$. Then $C_{k+1,n} \geqq C_{k,n}$. In the following argument we use the fact that a convergent series of non-negative terms may be rearranged at will.

$$\begin{aligned}\int_\Gamma f(\gamma)\, dm &= \lim_{k\to\infty} \sum_{n\in\mathcal{N}} C_{kn} = \sum_{k=1}^{\infty} \Big[\sum_{n\in\mathcal{N}} C_{k+1,n} - \sum_{n\in\mathcal{N}} C_{k,n}\Big] \\ &= \sum_k \sum_n (C_{k+1,n} - C_{k,n}) = \sum_n \sum_k (C_{k+1,n} - C_{k,n}) \\ &= \sum_n \lim_{k\to\infty} C_{k,n} = \sum_n \lim_{k\to\infty} \int_{P_n} f_k(\varrho_n)\, d\mu_n \\ &= \sum_{n\in\mathcal{N}} \int_{P_n} f(\varrho_n)\, d\mu_n = \sum_{n\in\mathcal{N}} \int_{P_n} f_n(\varrho_n)\, d\mu_n.\end{aligned} \tag{A 2.4}$$

6. Let $f_n(\varrho_n)$ be non-negative and measurable for each $n \in \mathcal{N}$. Then $f(\gamma) = f_n(\varrho_n)$ is non-negative and measurable. If $f_n(\varrho_n)$ is integrable for each $n \in \mathcal{N}$ and the series on the right side of (A 2.1) converges, then $f(\gamma)$ is integrable, and (A 2.1) is valid:

It can be shown that* there exists a non-decreasing sequence $f_1, f_2, \ldots$ of non-negative simple functions such that for every $\gamma \in \Gamma$, $\lim_{k\to\infty} f_k(\gamma) = f(\gamma)$. Then $\lim_{k\to\infty} f_k(\varrho_n) = f(\varrho_n) = f_n(\varrho_n)$. If $f_n(\varrho_n)$ is integrable, then

$$\lim_{k\to\infty} \int_{P_n} f_k(\varrho_n)\, d\mu_n = \int_{P_n} f_n(\varrho_n)\, d\mu_n. \tag{A 2.5}$$

Now if the series converges, we may reverse the steps in (A 2.4). By making use of Assertion 4, we see that $f(\gamma)$ is integrable and (A 2.1) is valid.

7. If $f(\gamma)$ is integrable on $(\Gamma, \mathcal{A}, m)$, then $f_n(\varrho_n) = f(\varrho_n)$ is integrable on $(P_n, \mathcal{L}_n, \mu_n)$ for each $n \in \mathcal{N}$, equation (A 2.1) holds, and the series converges absolutely:

If f is integrable, then by definition f^+ and f^- are integrable. Hence $f_n^+(\varrho_n)$ and $f_n^-(\varrho_n)$ are integrable for every $n \in \mathcal{N}$. By definition,

$$\int_\Gamma f\, dm = \int_\Gamma f^+\, dm - \int_\Gamma f^-\, dm = \sum_{n\in\mathcal{N}} \int_{P_n} f^+(\varrho_n)\, d\mu_n - \sum_{n\in\mathcal{N}} \int_{P_n} f_n^-(\varrho_n)\, d\mu_n. \tag{A 2.6}$$

Since both series converge, the right side of this equation converges absolutely and can be rearranged. Thus

$$\int_\Gamma f\, dm = \sum_{n\in\mathcal{N}} \Big[\int_{P_n} f_n^+(\varrho_n)\, d\mu_n - \int_{P_n} f_n^-(\varrho_n)\, d\mu_n\Big] = \sum_{n\in\mathcal{N}} \int_{P_n} f_n(\varrho_n)\, d\mu_n. \tag{A 2.7}$$

8. T_t is measure-preserving with respect to $(\Gamma, \mathcal{A}, m)$, for $t \geqq 0$:

For any $A \in \mathcal{A}$, $(T_t^{-1} A) \cap P_n = T_t^{-1}(A \cap P_n) = S_{nt}^{-1}(A \cap P_n) \in \mathcal{L}_n$ for every $n \in \mathcal{N}$. Hence $T_t^{-1} A \in \mathcal{A}$.

$$\begin{aligned} m(T_t^{-1} A) &= \sum_{n\in\mathcal{N}} \mu_n[(T_t^{-1} A) \cap P_n] = \sum_{n\in\mathcal{N}} \mu_n[S_{nt}^{-1}(A \cap P_n)] \\ &= \sum_{n\in\mathcal{N}} \mu_n(A \cap P_n) = m(A).\end{aligned} \tag{A 2.8}$$

---

* See [6], p. 155.

9. $T_{s+t} = T_s T_t$, for $t \geqq 0$, $s \geqq 0$:

This follows immediately from the same assertion for the S_{nt}.

10. $T_t \gamma$ is a measurable function of t and γ:

For each $n \in \mathscr{N}$; $S_{nt}\varrho_n$ is a measurable function of t and ϱ_n, *i.e.*, $S_{nt}\varrho_n$ is a measurable transformation from $(\mathscr{T} \times P_n, \mathscr{M} \times \mathscr{L}_n)$ into $(P_n, \mathscr{L}_n)$. Here $\mathscr{M}$ denotes the class of Lebesgue measurable sets in $\mathscr{T} = \{t \ni 0 \leqq t\}$. Thus for any $L_n \in \mathscr{L}_n \{(t, \varrho_n) \ni S_{nt}\varrho_n \in L_n\} \in \mathscr{M} \times \mathscr{L}_n$. Now, given any $A \in \mathscr{A}$, $\{(t, \gamma) \ni T_t \gamma \in A\} = \bigcup_{n \in \mathscr{N}} \{(t, \varrho_n) \ni T_t \varrho_n \in A\} = \bigcup_{n \in \mathscr{N}} G_n$, where $G_n = \{(t, \varrho_n) \ni S_{nt}\varrho_n \in A \cap P_n\}$. For each n, G_n is a set in the class $\mathscr{M} \times \mathscr{L}_n$. But $\mathscr{L}_n \subset \mathscr{A}$; hence $\mathscr{M} \times \mathscr{L}_n \subset \mathscr{M} \times \mathscr{A}$. Thus, for each n, $G_n \in \mathscr{M} \times \mathscr{A}$; hence $\{(t, \gamma) \ni T_t \gamma \in A\} \in \mathscr{M} \times \mathscr{A}$; *i.e.*, $T_t \gamma$ is a measurable transformation from $(\mathscr{T} \times \Gamma, \mathscr{M} \times \mathscr{A})$ into $(\Gamma, \mathscr{A})$.

11. $(\Gamma, \mathscr{A}, m, T_t)$ is a measure-preserving space.

This follows at once from 1, 2, 8, 9 and 10.

12. $g(\gamma)$ is invariant with respect to $(\Gamma, \mathscr{A}, m, T_t)$:

For each $t \geqq 0$, $g(T_t \gamma) = g_{(n)}(S_{nt}\varrho_n) = g_{(n)}(\varrho_n) = g(\gamma)$ except on a set $S \subset \Gamma$ such that $\mu_n(S \cap P_n) = 0$ for each $n \in \mathscr{N}$. Thus $m(S) = 0$.

13. $n(\gamma)$ is invariant with respect to $(\Gamma, \mathscr{A}, m, T_t)$:

$n(T_t \gamma) = n(S_{nt}\varrho_n) = n = n(\gamma)$.

14. $y(\gamma) = [n(\gamma), g(\gamma)]$ is a complete invariant vector on $(\Gamma, \mathscr{A}, m, T_t)$:

Let $h(\gamma)$ be any invariant function, *i.e.*, $h(T_t \gamma) = h(\gamma)$ a.e. $[m]$. For each fixed n, $h(S_{nt}\varrho_n) = h(\varrho_n)$ a.e. $[\mu_n]$. Hence, since $g_{(n)}$ is a complete invariant vector on $(P_n, \mathscr{L}_n, \mu_n, S_{nt})$, $h(\varrho_n) = H_n[g_{(n)}(\varrho_n)]$ a.e. $[\mu_n]$, where H_n is Borel-measurable on R^k. Let $z = (z_1, \ldots, z_\delta)$ be a point in R^δ, and set

$$H(z, g) = \begin{Bmatrix} H_n(g) & \text{if} & z = n \in \mathscr{N} \\ 0 & \text{if} & z \notin \mathscr{N} \end{Bmatrix}. \tag{A 2.9}$$

We shall show that H is Borel-measurable on $R^{k+\delta}$. If B is any Borel set in R^1 which does not contain the origin, then

$$\begin{aligned} H^{-1}(B) &= \{(z, g) \ni H(z, g) \in B\} = \bigcup_n \{(z, g) \ni H(z, g) \in B;\ z = n\} \\ &= \bigcup_n \{(z, g) \ni H_n(g) \in B;\ z = n\} = \bigcup_n \{(z, g) \ni g \in H_n^{-1}(B);\ z = n\} \\ &= \bigcup_n \{z = n\} \times H_n^{-1}(B). \end{aligned} \tag{A 2.10}$$

The set $\{z = n\}$ is just a point in R^δ and hence is a Borel set. $H_n^{-1}(B)$ is also a Borel set in R^k. Hence $H^{-1}(B)$ is a Borel set in $R^{k+\delta}$. Now let B_0 be the (Borel) set consisting of the origin in R^1.

$$H^{-1}(B_0) = \{(z, g) \ni z \notin \mathscr{N}\} \cup \bigcup_n \{(z, g) \ni H(z, g) = 0;\ z = n\}. \tag{A 2.11}$$

The set $\{(z, g) \ni z \notin \mathscr{N}\}$ is a Borel set in $R^{k+\delta}$, and for each n

$$\begin{aligned} \{(z, g) \ni H(z, g) = 0;\ z = n\} &= \{(z, g) \ni H_n(g) = 0;\ z = n\} \\ &= \{(z, g) \ni g \in H_n^{-1}(B_0);\ z = n\} = \{z = n\} \times H_n^{-1}(B_0). \end{aligned} \tag{A 2.12}$$

But this is again a Borel set; hence $H^{-1}(B_0)$ is a Borel set in $R^{k+\delta}$, and it follows that H is Borel-measurable on $R^{k+\delta}$. Now

$$h(\gamma) = h(\varrho_n) = H_n[g_{(n)}(\varrho_n)] = H[n, g_{(n)}(\varrho_n)] = H[n(\gamma), g(\gamma)], \tag{A 2.13}$$

except on a set such that $\mu_n(A \cap P_n) = 0$ for each $n \in \mathcal{N}$. Hence $h(\gamma) = H[n(\gamma), g(\gamma)]$ a.e. $[m]$.

15. $(\Gamma, \mathcal{A}, m, T_t, y)$ is a complete space.

This follows directly from 11 and 14.

References

[1] Grad, H.: Kinetic Theory and Statistical Mechanics (mimeographed notes). New York University 1950.

[2] Grad, H.: Statistical Mechanics, Thermodynamics, and Fluid Dynamics of Systems with an Arbitrary Number of Integrals. Comm. Pure Appl. Math. **5**, 455–494 (1952).

[3] Halmos, P. R.: Measure Theory. New York: D. Van Nostrand Co., Inc. 1950.

[4] Halmos, P. R.: Lectures on Ergodic Theory. The Mathematical Society of Japan 1956.

[5] Khinchin, A. I.: Mathematical Foundations of Statistical Mechanics. New York: Dover Publications, Inc. 1949.

[6] Munroe, M. E.: Introduction to Measure and Integration, Reading. Mass.: Addison-Wesley Publishing Co., Inc. 1953.

Institute of Mathematical Sciences
New York University
New York City

(Received April 11, 1960)

On the Minimality of Integrity Bases for Symmetric 3×3 Matrices

G. F. SMITH

Communicated by R. S. RIVLIN

1. Introduction

Finite integrity bases for five or fewer symmetric 3×3 matrices under the orthogonal transformation group have been derived by RIVLIN & SPENCER [*1*], [*4*]. There is no assurance given, however, that smaller sets of invariants will not also form integrity bases. In this paper, it will be shown that the integrity bases given in [*1*], [*4*] cannot be further reduced and they will be said to form minimal bases. In § 2, an expression is derived for $P_{i_1 i_2 \dots i_5}$, the number of linearly independent invariants of degree $i_1 i_2 \dots i_5$. A quantity of degree $i_1 i_2 \dots i_5$ is one of degree $i_1, i_2, \dots, i_5$ in the components of the matrices $\boldsymbol{A}_1, \boldsymbol{A}_2, \dots, \boldsymbol{A}_5$ respectively. The number of invariants $\vartheta_{i_1 i_2 \dots i_5}$ of degree $i_1 i_2 \dots i_5$ which are derivable from elements of the minimal basis of degree lower than $i_1 i_2 \dots i_5$ may be readily computed, and it is shown in § 3 that the number $P_{i_1 i_2 \dots i_5} - \vartheta_{i_1 i_2 \dots i_5}$ gives a lower bound for the number of invariants of degree $i_1 i_2 \dots i_5$ in the minimal basis. In § 4, the minimality of the bases given in [*1*], [*4*] is established by showing that the number of invariants of given degree in these bases coincides with the lower bound derived by the methods of § 2 and § 3.

2. The Number of Linearly Independent Invariants

In this section, an expression for the number $P_{i_1 i_2 \dots i_5}$ of linearly independent invariants of degree $i_1 i_2 \dots i_5$ will be derived. A function F of the matrices $\boldsymbol{A}_k$ $(=\|a_{ij}^{(k)}\|)$ is said to be invariant under the group of proper orthogonal transformations $\{\boldsymbol{R}\}$ if

$$F(a_{ij}^{(k)}) = F(\bar{a}_{ij}^{(k)}) \qquad (i,j=1,\dots,3) \tag{2.1}$$

where

$$\bar{a}_{ij}^{(k)} = l_{ip}\, l_{jq}\, a_{pq}^{(k)} \tag{2.2}$$

for all $\boldsymbol{l}$ $(=\|l_{ij}\|)$ such that

$$l_{ij}\, l_{ik} = \delta_{jk}, \qquad \det l_{ij} = +1\,. \tag{2.3}$$

It is convenient to use the notation

$$\begin{aligned} a_1^{(k)} &= a_{11}^{(k)}, & a_2^{(k)} &= a_{12}^{(k)}, & a_3^{(k)} &= a_{13}^{(k)}, \\ a_4^{(k)} &= a_{22}^{(k)}, & a_5^{(k)} &= a_{23}^{(k)}, & a_6^{(k)} &= a_{33}^{(k)}. \end{aligned} \tag{2.4}$$

Then, the invariance requirement (2.1) may be written as

$$F(a_j^{(k)}) = F(\bar{a}_j^{(k)}) \qquad (j = 1, \dots, 6). \tag{2.5}$$

The $\bar{a}_j^{(k)}$ corresponding to a given $\boldsymbol{l}$ are of course determined from (2.2) and (2.4) and may be presented in the form

$$\bar{a}_j^{(k)} = L_{ji}\, a_i^{(k)} \qquad (i, j = 1, \dots, 6) \tag{2.6}$$

where the matrices $\boldsymbol{L}$ $(= \|L_{ji}\|)$ are readily obtained from (2.2). Thus, the transformation matrix $\boldsymbol{L}$ of (2.6) corresponding to a rotation of ϑ degrees about the x_1 axis for which

$$\boldsymbol{l} = \left\| \begin{matrix} 1, & 0, & 0 \\ 0, & \cos\vartheta, & \sin\vartheta \\ 0, & -\sin\vartheta, & \cos\vartheta \end{matrix} \right\| \tag{2.7}$$

is given by

$$\boldsymbol{L} = \left\| \begin{matrix} 1, & 0, & 0, & 0, & 0, & 0 \\ 0, & \cos\vartheta, & \sin\vartheta, & 0, & 0, & 0 \\ 0, & -\sin\vartheta, & \cos\vartheta, & 0, & 0, & 0 \\ 0, & 0, & 0, & \cos^2\vartheta, & 2\cos\vartheta\sin\vartheta, & \sin^2\vartheta \\ 0, & 0, & 0, & -\cos\vartheta\sin\vartheta, & \cos^2\vartheta - \sin^2\vartheta, & \cos\vartheta\sin\vartheta \\ 0, & 0, & 0, & \sin^2\vartheta, & -2\cos\vartheta\sin\vartheta, & \cos^2\vartheta \end{matrix} \right\|. \tag{2.8}$$

Let $\boldsymbol{L}^n$ denote the n^{th} power of the matrix $\boldsymbol{L}$ and let S_n denote the trace of $\boldsymbol{L}^n$. Then, from (2.8),

$$S_1 = 2 + 2\cos\vartheta + 2\cos 2\vartheta$$

and

$$S_n = 2 + 2\cos n\vartheta + 2\cos 2n\vartheta. \tag{2.9}$$

The number P_1 of linearly independent invariants of degree one in the components $a_j^{(1)}$ of a single matrix $\boldsymbol{A}_1$ is obtained by averaging over $\{\boldsymbol{R}\}$ the trace of the transformation matrix $\boldsymbol{L}$ of the independent components $a_j^{(1)}$. This procedure is accomplished by integration over the group manifold and the result is given by

$$P_1 = \frac{1}{2\pi} \int_0^{2\pi} \operatorname{Tr} \boldsymbol{L}(\vartheta)\, (1 - \cos\vartheta)\, d\vartheta \tag{2.10}$$

where $\operatorname{Tr} \boldsymbol{L}(\vartheta) = S_1$ is given by (2.9). Details may be found in the standard references on group representation theory [*2*], [*3*], or in the papers of I. Schur [*5*].

In order to determine the number $P_{i_1 i_2 \dots i_5}$ of linearly independent invariants of degree $i_1\, i_2 \dots i_5$, it is necessary to average over $\{\boldsymbol{R}\}$ the trace of the transformation matrix of the independent quantities

$$a_{j_1}^{(1)} a_{j_2}^{(1)} \dots a_{j_{i_1}}^{(1)} a_{k_1}^{(2)} a_{k_2}^{(2)} \dots a_{k_{i_2}}^{(2)} \dots a_{p_1}^{(5)} a_{p_2}^{(5)} \dots a_{p_{i_5}}^{(5)} \tag{2.11}$$

where

$$j_1, j_2, \dots, p_{i_5} = 1, \dots, 6$$

and

$$j_1 \leqq j_2 \leqq \dots \leqq j_{i_1}, \quad k_1 \leqq k_2 \leqq \dots \leqq k_{i_2}, \dots, p_1 \leqq p_2 \leqq \dots \leqq p_{i_5}.$$

Let the matrix $\boldsymbol{L}^{[n]}$ $(=\boldsymbol{L}\times\boldsymbol{L}\times\cdots\times\boldsymbol{L})$ denote the Kronecker* n^{th} power of the matrix $\boldsymbol{L}$ and the matrix $\boldsymbol{L}^{(n)}$ the symmetrized Kronecker n^{th} power of the matrix $\boldsymbol{L}$. Then, if the transformation properties of each of the sets of quantities $a_j^{(k)}$ $(j=1,\dots,6;\ k=1,\dots,5)$ are given by the matrix $\boldsymbol{L}$, the transformation properties of the 6^n quantities $a_{j_1}^{(1)}\, a_{j_2}^{(2)}\dots a_{j_n}^{(n}$ $(j_1,\dots,j_n=1,\dots,6)$ are given by the matrix $\boldsymbol{L}^{[n]}$ and the transformation properties of the $\binom{6+n-1}{n}$ quantities

$$a_{j_1}^{(1)}\, a_{j_2}^{(1)}\dots a_{j_n}^{(1)} \qquad (j_1,\dots,j_n=1,\dots,6;\ j_1\leqq j_2\leqq\cdots\leqq j_n)$$

are given by the matrix $\boldsymbol{L}^{(n)}$.

It is readily seen that the transformation matrix for the

$$\binom{6+i_1-1}{i_1}\cdot\binom{6+i_2-1}{i_2}\cdot\ \cdots\ \binom{6+i_5-1}{i_5}$$

quantities (2.11) is given by

$$\boldsymbol{L}^{(i_1)}\times\boldsymbol{L}^{(i_2)}\times\cdots\times\boldsymbol{L}^{(i_5)} \tag{2.12}$$

when the transformation matrix for the quantities $a_j^{(k)}$ $(j=1,\dots,6;\ k=1,\dots,5)$ is given by $\boldsymbol{L}$. The matrix (2.12) is the Kronecker product of the matrices $\boldsymbol{L}^{(i_1)},\boldsymbol{L}^{(i_2)},\dots,\boldsymbol{L}^{(i_5)}$ which are in turn the symmetrized Kronecker $i_1, i_2,\dots,i_5$ powers of the matrix $\boldsymbol{L}$.

It is easily shown [2] that the trace of the Kronecker product of a number of matrices is equal to the product of the traces of the individual matrices. Thus,

$$\operatorname{Tr}(\boldsymbol{L}^{(i_1)}\times\boldsymbol{L}^{(i_2)}\times\cdots\times\boldsymbol{L}^{(i_5)})=\operatorname{Tr}\boldsymbol{L}^{(i_1)}\operatorname{Tr}\boldsymbol{L}^{(i_2)}\dots\operatorname{Tr}\boldsymbol{L}^{(i_5)}. \tag{2.13}$$

The trace of the symmetrized Kronecker product $\boldsymbol{L}^{(n)}$ is given in terms of the traces $S_1, S_2,\dots,S_n$ of the matrix products $\boldsymbol{L}^1,\boldsymbol{L}^2,\dots,\boldsymbol{L}^n$ by the relation (see [2], p. 106)

$$\operatorname{Tr}\boldsymbol{L}^{(n)}=\sum\frac{1}{\alpha_1!\,\alpha_2!\dots\alpha_n!}\left(\frac{S_1}{1}\right)^{\alpha_1}\left(\frac{S_2}{2}\right)^{\alpha_2}\cdots\left(\frac{S_n}{n}\right)^{\alpha_n} \tag{2.14}$$

where the summation is over the set of all positive integers $\alpha_1,\alpha_2,\dots,\alpha_n$ such that

$$\alpha_1+2\alpha_2+\cdots+n\,\alpha_n=n.$$

Thus,

$$\begin{aligned}
\operatorname{Tr}\boldsymbol{L}\ \ &=S_1,\\
\operatorname{Tr}\boldsymbol{L}^{(2)}&=\frac{1}{2}(S_1^2+S_2),\\
\operatorname{Tr}\boldsymbol{L}^{(3)}&=\frac{1}{3!}(S_1^3+3\,S_1\,S_2+2\,S_3),\\
\operatorname{Tr}\boldsymbol{L}^{(4)}&=\frac{1}{4!}(S_1^4+6\,S_1^2\,S_2+8\,S_1\,S_3+3\,S_2^2+6\,S_4),\\
\operatorname{Tr}\boldsymbol{L}^{(5)}&=\frac{1}{5!}(S_1^5+10\,S_1^3\,S_2+20\,S_1^2\,S_3+15\,S_1\,S_2^2+30\,S_1\,S_4+20\,S_2\,S_3+24\,S_5).
\end{aligned} \tag{2.15}$$

The number of linearly independent invariants of degree $i_1\, i_2\dots i_5$ is then given by $P_{i_1 i_2\dots i_5}$ where

$$P_{i_1 i_2\dots i_5}=\frac{1}{2\pi}\int_0^{2\pi}\operatorname{Tr}\boldsymbol{L}^{(i_1)}\operatorname{Tr}\boldsymbol{L}^{(i_2)}\dots\operatorname{Tr}\boldsymbol{L}^{(i_5)}(1-\cos\vartheta)\,d\vartheta \tag{2.16}$$

* For the definition and properties of the Kronecker products, see [2], Chap. 4.

where $\mathrm{Tr}\,\boldsymbol{L}^{(i_1)}$, $\mathrm{Tr}\,\boldsymbol{L}^{(i_2)}$, ..., $\mathrm{Tr}\,\boldsymbol{L}^{(i_5)}$ are functions of ϑ which are given by (2.9) and (2.14). With (2.14) and (2.15), the quantity $\mathrm{Tr}\,\boldsymbol{L}^{(i_1)}\,\mathrm{Tr}\,\boldsymbol{L}^{(i_2)}\ldots\mathrm{Tr}\,\boldsymbol{L}^{(i_5)}$ is expressible as a polynomial in the quantities $S_1, S_2, \ldots, S_n$. Thus, $P_{i_1 i_2 \ldots i_5}$ may be evaluated when the quantities

$$s_1^i\, s_2^j \ldots s_n^k = \frac{1}{2\pi} \int_0^{2\pi} S_1^i\, S_2^j \ldots S_n^k (1-\cos\vartheta)\, d\vartheta \tag{2.17}$$

have been tabulated. Table 1 gives the values of the $s_1^i\, s_2^j \ldots s_n^k$ which are required in this study.

Table 1

$s_1=1$	$s_1^2=2$	$s_1^3=5$	$s_1^4=16$	$s_1^5=62$	$s_1^6=272$
	$s_2=2$	$s_1 s_2=3$	$s_1^2 s_2=6$	$s_1^3 s_2=14$	$s_1^4 s_2=42$
		$s_3=2$	$s_2^2=8$	$s_1 s_2^2=14$	$s_1^3 s_3=5$
					$s_1^2 s_2^2=28$

3. The Number of Basis Elements

In this section, an expression for the lower bound on the number of elements of the minimal basis of degree $i_1\, i_2 \ldots i_5$ will be derived. Let

$\gamma_{i_1 i_2 \ldots i_5}$ = the number of elements of degree $i_1\, i_2 \ldots i_5$ in the minimal basis $J_1, \ldots, J_m$.

$\beta_{i_1 i_2 \ldots i_5}$ = the number of elements of degree $i_1\, i_2 \ldots i_5$ in the basis $I_1, \ldots, I_n$ given in [1], [4].

$P_{i_1 i_2 \ldots i_5}$ = the number of linearly independent invariants of degree $i_1\, i_2 \ldots i_5$.

$\vartheta_{i_1 i_2 \ldots i_5}$ = the number of invariants of degree $i_1\, i_2 \ldots i_5$ which may be formed from invariants of degree lower than $i_1\, i_2 \ldots i_5$ in the minimal basis $J_1, \ldots, J_m$.

It is clear that $\beta_{i_1 i_2 \ldots i_5} \geqq \gamma_{i_1 i_2 \ldots i_5}$ since the $\gamma_{i_1 i_2 \ldots i_5}$ represent the number of elements of degree $i_1\, i_2 \ldots i_5$ in the minimal basis. The quantities $P_{i_1 i_2 \ldots i_5} - \vartheta_{i_1 i_2 \ldots i_5}$ represent the number of invariants of degree $i_1\, i_2 \ldots i_5$ in the minimal basis provided that the $\vartheta_{i_1 i_2 \ldots i_5}$ invariants of degree $i_1\, i_2 \ldots i_5$ which are formed from the elements of the minimal basis of degree lower than $i_1\, i_2 \ldots i_5$ are all linearly independent. Since this is not necessarily the case, $\vartheta_{i_1 i_2 \ldots i_5}$ will give the upper bound for the number of linearly independent invariants of degree $i_1\, i_2 \ldots i_5$ which may be obtained from the products of elements of degree lower than $i_1\, i_2 \ldots i_5$ in the minimal basis. Consequently, the quantity $P_{i_1 i_2 \ldots i_5} - \vartheta_{i_1 i_2 \ldots i_5}$ represents a lower bound for the number of elements of the minimal basis of degree $i_1\, i_2 \ldots i_5$. Thus,

$$P_{i_1 i_2 \ldots i_5} - \vartheta_{i_1 i_2 \ldots i_5} \leqq \gamma_{i_1 i_2 \ldots i_5} \leqq \beta_{i_1 i_2 \ldots i_5}. \tag{3.1}$$

The quantities $\vartheta_{i_1 i_2 \ldots i_5}$ may be computed from the expression

$$\vartheta_{ij\ldots k} = \sum \binom{\gamma_{i_1 j_1 \ldots k_1} + n_1 - 1}{n_1} \binom{\gamma_{i_2 j_2 \ldots k_2} + n_2 - 1}{n_2} \cdots \binom{\gamma_{i_p j_p \ldots k_p} + n_p - 1}{n_p} \tag{3.2}$$

where the summation is over the distinct sets of positive integers $(n_r, i_r, j_r, \ldots, k_r)$ such that

$$\begin{aligned} n_1 i_1 + n_2 i_2 + \cdots + n_p i_p &= i, \\ n_1 j_1 + n_2 j_2 + \cdots + n_p j_p &= j, \\ &\vdots \\ n_1 k_1 + n_2 k_2 + \cdots + n_p k_p &= k, \end{aligned}$$

and where at least one of the inequalities

$$i_r \leqq i, \quad j_r \leqq j, \ldots, k_r \leqq k \quad (r = 1, \ldots, p)$$

must be a strict inequality.

4. Minimality of the Integrity Bases

In this section, the minimality of the integrity bases derived by RIVLIN & SPENCER [1], [4] for five or fewer symmetric 3×3 matrices will be established. These integrity bases are given in terms of the quantities (4.1), ..., (4.5) below.

I. Invariants involving a single matrix:

$$\operatorname{Tr} \boldsymbol{A}_i, \ \operatorname{Tr} \boldsymbol{A}_i^2, \ \operatorname{Tr} \boldsymbol{A}_i^3. \tag{4.1}$$

II. Invariants involving two matrices:

$$\operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j, \ \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j^2, \ \operatorname{Tr} \boldsymbol{A}_i^2 \boldsymbol{A}_j, \ \operatorname{Tr} \boldsymbol{A}_i^2 \boldsymbol{A}_j^2 \quad (i<j). \tag{4.2}$$

III. Invariants involving three matrices:

$$\left.\begin{aligned} &\text{(i)} \quad \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j \boldsymbol{A}_k \quad (i<j<k). \\ &\text{(ii)} \quad \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j \boldsymbol{A}_k^2, \ \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j^2 \boldsymbol{A}_k^2 \quad (i<j<k) \end{aligned}\right\} \tag{4.3}$$

and invariants obtained from (ii) by permuting the matrices $\boldsymbol{A}_i, \ldots, \boldsymbol{A}_k$ cyclically.

IV. Invariants involving four matrices:

$$\begin{aligned} &\text{(i)} \quad \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j \boldsymbol{A}_k \boldsymbol{A}_l, \quad \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j \boldsymbol{A}_l \boldsymbol{A}_k \quad (i<j<k<l). \\ &\text{(ii)} \quad \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j \boldsymbol{A}_k \boldsymbol{A}_l^2, \quad \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_k \boldsymbol{A}_j \boldsymbol{A}_l^2 \quad (i<j<k<l) \end{aligned}$$

and invariants obtained from (ii) by permuting the matrices $\boldsymbol{A}_i, \ldots, \boldsymbol{A}_l$ cyclically.

$$\begin{aligned} \text{(iii)} \quad &\operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j \boldsymbol{A}_k^2 \boldsymbol{A}_l^2, \quad \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_k \boldsymbol{A}_j^2 \boldsymbol{A}_l^2, \quad \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_l \boldsymbol{A}_j^2 \boldsymbol{A}_k^2, \quad \operatorname{Tr} \boldsymbol{A}_j \boldsymbol{A}_k \boldsymbol{A}_i^2 \boldsymbol{A}_l^2, \\ &\operatorname{Tr} \boldsymbol{A}_j \boldsymbol{A}_l \boldsymbol{A}_i^2 \boldsymbol{A}_k^2, \quad \operatorname{Tr} \boldsymbol{A}_k \boldsymbol{A}_l \boldsymbol{A}_i^2 \boldsymbol{A}_j^2 \quad (i<j<k<l). \end{aligned}$$

$$\text{(iv)} \quad \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_l \boldsymbol{A}_j \boldsymbol{A}_k \boldsymbol{A}_l^2 \quad (i<j<k<l) \tag{4.4}$$

and invariants obtained from (iv) by permuting the matrices $\boldsymbol{A}_i, \ldots, \boldsymbol{A}_l$ cyclically

V. Invariants involving five matrices:

$$\left.\begin{aligned} \text{(i)} \quad &\operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j \boldsymbol{A}_k \boldsymbol{A}_l \boldsymbol{A}_m, \ \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j \boldsymbol{A}_l \boldsymbol{A}_m \boldsymbol{A}_k, \ \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j \boldsymbol{A}_m \boldsymbol{A}_k \boldsymbol{A}_l, \\ &\operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_k \boldsymbol{A}_l \boldsymbol{A}_j \boldsymbol{A}_m, \ \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_k \boldsymbol{A}_j \boldsymbol{A}_m \boldsymbol{A}_l, \ \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_l \boldsymbol{A}_j \boldsymbol{A}_k \boldsymbol{A}_m. \end{aligned} \right\} (i<j<k<l<m)$$

$$\left.\begin{aligned} \text{(ii)} \quad &\operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j \boldsymbol{A}_k \boldsymbol{A}_l \boldsymbol{A}_m^2, \ \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_j \boldsymbol{A}_l \boldsymbol{A}_k \boldsymbol{A}_m^2, \ \operatorname{Tr} \boldsymbol{A}_i \boldsymbol{A}_l \boldsymbol{A}_k \boldsymbol{A}_j \boldsymbol{A}_m^2, \\ &\operatorname{Tr} \boldsymbol{A}_j \boldsymbol{A}_i \boldsymbol{A}_k \boldsymbol{A}_l \boldsymbol{A}_m^2 \quad (i<j<k<l<m) \end{aligned}\right\} \tag{4.5}$$

and invariants obtained from (ii) by permuting the matrices $\boldsymbol{A}_i, \ldots, \boldsymbol{A}_m$ cyclically.

The minimality of the bases given in [*1*], [*4*] will be established by showing that the number of elements $\beta_{i_1 i_2 \ldots i_5}$ of degree $i_1\, i_2 \ldots i_5$ in these bases coincides with the number of elements $\gamma_{i_1 i_2 \ldots i_5}$ of degree $i_1\, i_2 \ldots i_5$ in the minimal bases. It is evident that the number of elements $\gamma_{i_1 i_2 \ldots i_5}$ of degree $i_1\, i_2 \ldots i_5$ in the minimal basis is equal to the number of elements $\gamma_{j_1 j_2 \ldots j_5}$ of degree $j_1\, j_2 \ldots j_5$ where $j_1\, j_2 \ldots j_5$ is any permutation of $i_1\, i_2 \ldots i_5$. Since the number of elements of the bases in [*1*], [*4*] have the same property, *i.e.*, $\beta_{i_1 i_2 \ldots i_5} = \beta_{j_1 j_2 \ldots j_5}$ for $j_1\, j_2 \ldots j_5$ any permutation of $i_1\, i_2 \ldots i_5$, it will be necessary in the following to establish that $\gamma_{i_1 i_2 \ldots i_5} = \beta_{i_1 i_2 \ldots i_5}$ for only one of the permutations of $i_1\, i_2 \ldots i_5$.

Single matrix $\boldsymbol{A}_1$. The integrity basis is given by the quantities (4.1) with $i=1$. Thus,

$$\beta_1 = \beta_2 = \beta_3 = 1 .$$

From (2.16) and Table 1,

$$P_1 = 1 , \qquad P_2 = 2 , \qquad P_3 = 3 .$$

Then, from (3.1),

$$P_1 - \vartheta_1 = 1 \leqq \gamma_1 \leqq \beta_1 = 1 \Rightarrow \gamma_1 = \beta_1 = 1 ,$$

$$P_2 - \vartheta_2 = 1 \leqq \gamma_2 \leqq \beta_2 = 1 \Rightarrow \gamma_2 = \beta_2 = 1 ,$$

$$P_3 - \vartheta_3 = 1 \leqq \gamma_3 \leqq \beta_3 = 1 \Rightarrow \gamma_3 = \beta_3 = 1$$

where, from (3.2),

$$\vartheta_1 = 0 , \qquad \vartheta_2 = 1 , \qquad \vartheta_3 = 2 .$$

Since $\beta_n = 0$ implies that $\gamma_n = 0$, it is seen that $\beta_i = \gamma_i$ and hence the quantities (4.1) with $i=1$ form a minimal basis for a single matrix $\boldsymbol{A}_1$.

Two matrices $\boldsymbol{A}_1$, $\boldsymbol{A}_2$. The integrity basis is given by the quantities (4.1) and (4.2) with $i, j = 1, 2$. It is clear from the previous section that the quantities (4.1) with $i = 1, 2$ belong to the minimal basis and it is only necessary to verify that the quantities (4.2) with $i, j = 1,2$ also belong to the minimal basis. From (4.2),

$$\beta_{11} = \beta_{12} = \beta_{22} = 1 .$$

From (2.16) and Table 1,

$$P_{11} = 2 , \qquad P_{12} = 4 , \qquad P_{22} = 9 .$$

Then, from (3.1),

$$P_{11} - \vartheta_{11} = 1 \leqq \gamma_{11} \leqq \beta_{11} = 1 \Rightarrow \gamma_{11} = \beta_{11} = 1 ,$$

$$P_{12} - \vartheta_{12} = 1 \leqq \gamma_{12} \leqq \beta_{12} = 1 \Rightarrow \gamma_{12} = \beta_{12} = 1 ,$$

$$P_{22} - \vartheta_{22} = 1 \leqq \gamma_{22} \leqq \beta_{22} = 1 \Rightarrow \gamma_{22} = \beta_{22} = 1$$

where, from (3.2),

$$\vartheta_{11} = 1 , \qquad \vartheta_{12} = 3 , \qquad \vartheta_{22} = 8 .$$

Hence, the quantities (4.1) and (4.2) with $i, j = 1,2$ form a minimal basis for two matrices $\boldsymbol{A}_1, \boldsymbol{A}_2$.

Three matrices $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3$. The integrity basis is given by the quantities (4.1) (4.2) and (4.3) with $i, j, k = 1, 2, 3$. It is seen from the previous sections that the quantities (4.1) and (4.2) with $i, j = 1, 2, 3$ belong to the minimal basis, and it is only necessary to verify that the quantities (4.3) with $i, j, k = 1, 2, 3$ belong

to the minimal basis. From (4.3),

$$\beta_{111}=1\,,\quad \beta_{112}=1\,,\quad \beta_{122}=1\,.$$

From (2.16) and Table 1,

$$P_{111}=5\,,\quad P_{112}=11\,,\quad P_{122}=26\,.$$

Then, from (3.1),

$$P_{111}-\vartheta_{111}=1\leqq\gamma_{111}\leqq\beta_{111}=1\Rightarrow\gamma_{111}=\beta_{111}=1\,,$$
$$P_{112}-\vartheta_{112}=1\leqq\gamma_{112}\leqq\beta_{112}=1\Rightarrow\gamma_{112}=\beta_{112}=1\,,$$
$$P_{122}-\vartheta_{122}=1\leqq\gamma_{122}\leqq\beta_{122}=1\Rightarrow\gamma_{122}=\beta_{122}=1$$

where, from (3.2),

$$\vartheta_{111}=4\,,\quad \vartheta_{112}=10\,,\quad \vartheta_{122}=25\,.$$

Hence, the quantities (4.1), (4.2) and (4.3) with $i, j, k=1, 2, 3$ form a minimal basis for three matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_3$.

Four matrices $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3, \boldsymbol{A}_4$. The integrity basis is given by the quantities (4.1), (4.2), (4.3) and (4.4) with $i, j, k, l=1, 2, 3, 4$. It is seen from the previous sections that the quantities (4.1), (4.2) and (4.3) with $i, j, k=1, 2, 3, 4$ belong to the minimal basis, and it is only necessary to verify that the quantities (4.4) with $i, j, k, l=1, 2, 3, 4$ belong to the minimal basis. From (4.4),

$$\beta_{1111}=2\,,\quad \beta_{1112}=2\,,\quad \beta_{1122}=1\,,\quad \beta_{1113}=1\,.$$

From (2.16) and Table 1,

$$P_{1111}=16\,,\quad P_{1112}=38\,,\quad P_{1122}=96\,,\quad P_{1113}=68\,.$$

Then, from (3.1),

$$P_{1111}-\vartheta_{1111}=2\leqq\gamma_{1111}\leqq\beta_{1111}=2\Rightarrow\gamma_{1111}=\beta_{1111}=2\,,$$
$$P_{1112}-\vartheta_{1112}=2\leqq\gamma_{1112}\leqq\beta_{1112}=2\Rightarrow\gamma_{1112}=\beta_{1112}=2\,,$$
$$P_{1122}-\vartheta_{1122}=1\leqq\gamma_{1122}\leqq\beta_{1122}=1\Rightarrow\gamma_{1122}=\beta_{1122}=1\,,$$
$$P_{1113}-\vartheta_{1113}=1\leqq\gamma_{1113}\leqq\beta_{1113}=1\Rightarrow\gamma_{1113}=\beta_{1113}=1$$

where, from (3.2),

$$\vartheta_{1111}=14\,,\quad \vartheta_{1112}=36\,,\quad \vartheta_{1122}=95\,,\quad \vartheta_{1113}=67\,.$$

Hence, the quantities (4.1), (4.2), (4.3) and (4.4) with $i, j, k, l=1, 2, 3, 4$ form a minimal basis for four matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_4$.

Five matrices $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3, \boldsymbol{A}_4, \boldsymbol{A}_5$. The integrity basis is given by the quantities (4.1), (4.2), (4.3), (4.4) and (4.5) with $i, j, k, l, m=1, 2, 3, 4, 5$. It is seen from the previous sections that the quantities (4.1), (4.2), (4.3) and (4.4) with $i, j, k, l=1, 2, 3, 4, 5$ belong to the minimal basis and it is only necessary to verify that the quantities (4.5) with $i, j, k, l, m=1, 2, 3, 4, 5$ belong to the minimal basis. From (4.5),

$$\beta_{11111}=6\,,\quad \beta_{11112}=4\,.$$

From (2.16) and Table 1,

$$P_{11111}=62\,,\quad P_{11112}=157\,.$$

Then, from (3.1),

$$P_{11111} - \vartheta_{11111} = 6 \leqq \gamma_{11111} \leqq \beta_{11111} = 6 \Rightarrow \gamma_{11111} = \beta_{11111} = 6,$$
$$P_{11112} - \vartheta_{11112} = 4 \leqq \gamma_{11112} \leqq \beta_{11112} = 4 \Rightarrow \gamma_{11112} = \beta_{11112} = 4$$

where, from (3.2),

$$\vartheta_{11111} = 56, \quad \vartheta_{11112} = 153.$$

Hence, the quantities (4.1), ..., (4.5) with $i, j, k, l, m = 1, 2, 3, 4, 5$ form a minimal basis for five matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_5$.

References

[1] Spencer, A. J. M., & R. S. Rivlin: Arch. Rational Mech. Anal. **2**, 435—446 (1959).
[2] Murnaghan, F. D.: The Theory of Group Representations. Baltimore 1938.
[3] Wigner, E. P.: Group Theory. Academic Press 1959.
[4] Spencer, A. J. M., & R. S. Rivlin: Arch. Rational Mech. Anal. **4**, 214—230 (1960).
[5] Schur, I.: Sitzungsber. Preuß. Akad. 189—208, 297—321, 346—355 (1924).

Lehigh University
Bethlehem, Pennsylvania

(Received March 19, 1960)

Fourier-Transform Methods in the Theory of Approximation

P. L. BUTZER

Communicated by C. MÜLLER

§ 1. Introduction and the Transform Method[1]

Integral transform methods such as Fourier-transform methods have proven to be of great importance in the solution of initial and boundary value problems in mathematical physics. The general method is to transform a given (partial) differential equation, involving an unknown function, into an equation involving the Fourier transform (or some similar transform) of the original function. This new equation is then solved, and an appropriate inversion theorem is applied to obtain the solution in terms of the original function space. These methods have been investigated not only from a formal but also from a rigorous point of view by S. BOCHNER [*4*], G. DOETSCH [*21*], J. L. B. COOPER [*18*], E. C. TITCHMARSH [*33*] and many others.

The object of this paper is to study such Fourier-transform methods in connection with a class of problems in the theory of approximation. This method was first considered by the author in [*11*], and it is the aim here to continue these investigations. Some of these and related problems have been investigated from an entirely different point of view by E. HILLE ([*27*], p. 323 and p. 386) and the author [*8, 9, 10*] using a semi-group method. The latter method has the same degree of generality as the Fourier-transform method, both possessing a general approach and unified theory. It may be remarked that in case the functions under consideration belong to non-reflexive Banach spaces, *e.g.* $L_1(-\infty, \infty)$, or the approximating singular integrals do not satisfy the semi-group property, the Fourier method is generally the more advantageous of the two.

We proceed with an outline of the detailed results and the transform method.

Let us denote the Fourier transform of the function $f \in L_1(-\infty, \infty)$ by

$$(1.1) \qquad \mathfrak{F}(f) = \hat{f}, \quad \hat{f}(v) = \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{\infty} f(x)\, e^{-ivx}\, dx,$$

and the Fourier-Stieltjes transform of g by

$$(1.2) \qquad \mathfrak{F}\mathfrak{S}(g) = \check{g}, \quad \check{g}(v) = \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{\infty} e^{-ivx}\, dg(x),$$

g being of bounded variation on $(-\infty, \infty)$.

[1] The results of this paper were presented at a meeting of the DMV in Münster, October 20, 1959. A summary of these results appeared in the Comptes rendus de l'Académie des Sciences (Paris); see [*12*].

Given a general singular integral

$$(1.3)\qquad J_\varrho(x) = \frac{\varrho}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x+u)\, k(\varrho u)\, du$$

with parameter $\varrho > 0$ and kernel $k(u)$ (see §3 for properties of k), the problem is to consider the convergence of $J_\varrho(x)$ towards $f(x)$ for large ϱ and to characterize the class of functions f for which the degree of approximation of f by $J_\varrho(x)$ in the $L_1(-\infty, \infty)$ norm is exactly of order $O(\varrho^{-\gamma})$, $\gamma \geqq 0$, *i.e.* for which

$$(1.4)\qquad D_\varrho(J; f) = \|f(x) - J_\varrho(x)\| = O(\varrho^{-\gamma}), \qquad \varrho \uparrow \infty.$$

The Fourier-transform method in handling problems of this type may be formally explained as follows. Let us suppose that the degree of approximation of f by $J_\varrho(x)$ is given by

$$D_\varrho(J; f) = o(\varrho^{-\gamma}), \qquad \varrho \uparrow \infty.$$

An application of Fourier transforms to the latter equation gives

$$|\hat{f}(v)\,[1 - \hat{k}(v/\varrho)]| \leqq \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{\infty} |[f(x) - J_\varrho(x)]\, e^{-ivx}|\, dx = o(\varrho^{-\gamma}),$$

or

$$(1.5)\qquad \lim_{\varrho \uparrow \infty} \varrho^\gamma \hat{f}(v)\,[1 - \hat{k}(v/\varrho)] = 0, \qquad \text{for all } v.$$

For certain kernels of frequent use in mathematical physics (see §3), the transformed equation

$$c\,|v|^\gamma \hat{f}(v) = 0, \qquad \text{for all } v,$$

for constant $c > 0$, results. Then a suitable inversion theorem is applied, yielding the solution $f = 0$ a.e. in $(-\infty, \infty)$.

In the case

$$(1.6)\qquad D_\varrho(J; f) = O(\varrho^{-\gamma}),$$

this equation is transformed via Fourier transforms (for suitable kernels) into

$$(1.7)\qquad c\,|v|^\gamma \hat{f}(v) = \hat{g}(v), \qquad \text{for all } v,$$

where g is of bounded variation on $(-\infty, \infty)$. In terms of the original functions, the solution to our problem then is that f belongs to a certain class $\boldsymbol{K}$. In the case $\gamma = 2$, the class $\boldsymbol{K}$ is the class of functions for which the derivative $f'(x)$ is of bounded variation on $(-\infty, \infty)$.

Problems of this type are classified under the *inverse theorems* in the theory of approximation. The corresponding *direct problem* is the following: Given an integral of type (1.3) with an appropriate kernel and given that f belongs to a certain class $\boldsymbol{K}$ such that a condition of type (1.7) is satisfied, then one must prove that (1.6) holds. To solve this latter problem, the transform method may again be applied.

The problems formulated above must be approached from a rigorous point of view. This gives rise to the following definition:

Let $\varphi(\varrho)$ be a positive non-increasing function of ϱ such that $\varphi(\varrho) \downarrow 0$ as $\varrho \uparrow \infty$. If there exists a class $\boldsymbol{K}$ of functions f such that

(i) $D_\varrho(J; f) = o(\varphi(\varrho))$ implies that $f(x)$ is constant,

(ii) $D_\varrho(J; f) = O(\varphi(\varrho))$ implies that $f \in \boldsymbol{K}$,

(iii) for any $f \in \boldsymbol{K}$ one has $D_\varrho(J; f) = O(\varphi(\varrho))$,

then the singular integral $J_\varrho(x)$ is said to be *saturated* with order $O(\varphi(\varrho))$, and $\boldsymbol{K}$ is called the *saturation class* of the given $J_\varrho(x)$.

This concept was first introduced into the theory of approximation by J. FAVARD [23] and formulated for semi-group operators by the writer [8]. It enables one to restate the above problems in accurate terms.

The method of approach broadly outlined above is not restricted to (non-periodic) functions of class $L(-\infty, \infty)$. In the case of periodic functions f, having period 2π, one may consider integrals of the form

$$J_\varrho(x) = \frac{\varrho}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x+u)\, k(\varrho u)\, du,$$

and in discussing problems of the type under consideration, instead of applying Fourier transforms, one may employ the Fourier coefficients as a suitable transform:

$$\hat{f}_F(n) = \left(1/\sqrt{2\pi}\right) \int_{-\pi}^{\pi} f(x)\, e^{-inx}\, dx \qquad (n = 0, \pm 1, \pm 2, \ldots).$$

This *general* method is implicitly contained in [11], especially Theorems 5.1 and 6.3, and is being investigated at present by G. SUNOUCHI & C. WATARI [32]. But here it must be pointed out that the saturation classes for various *special* singular integrals of functions $f \in C[-\pi, +\pi]$ or $L_1(-\pi, +\pi)$, which can be interpreted as methods of summation of Fourier series of f, have been recently investigated by various authors. This is the case with the Cesàro (C, k), Riesz $R(\lambda, k)$, Hölder (H, k), Abel and Lambert methods of summation, which have been treated by S. ALJANČIĆ [2], H. BUCHWALTER [7], J. FAVARD [24], A. H. TURECKIĬ [34] and M. ZAMANSKY [35]. But the methods of proof in these papers are often restricted to one special singular integral and cannot be generalized to a wider class of such integrals.

In case the functions f are non-periodic and defined over a fixed finite interval (which by a linear transformation may be reduced to the interval $[-1, +1]$), the role played above by the Fourier transform may, for integrals of the type (1.3) but defined over $[-1, +1]$, be taken over by the finite Legendre transform

$$\hat{f}_L(n) = \int_{-1}^{1} f(x)\, L_n(x)\, dx,$$

where $L_n(x)$ denotes the Legendre polynomial of degree n.

In the resolution of boundary value problems by operational methods in the case when the variables range over a finite interval, G. DOETSCH [20] first suggested the use of finite sine or cosine transforms. These methods have been extended by I. N. SNEDDON [31] and especially by R. V. CHURCHILL [16].

In §2 of the present paper we quote several known results which are somewhat scattered throughout the literature and are needed in our work. In §3 we state the conditions the kernel k in question may have to possess, and we demonstrate two direct approximation theorems, the proofs of which are free of transform methods. In §4 two inverse theorems are established, the proofs depending upon the Fourier-transform methods. In §5, the Fourier method is applied to direct theorems of a type similar to those of §3. There is also a discussion on the saturation class of the singular integral (1.3). In § 6 the problems treated in the previous sections are extended to $L_p(-\infty, \infty)$ spaces for $1<p\leqq 2$, which are reflexive. In § 7 we discuss applications of the general theorems presented to Fejér's singular integral and that of Poisson-Cauchy as well as Gauss-Weierstrass. In § 8 the saturation classes for Picard's and Jackson's singular integrals are treated, both integrals not being semi-group operators. Finally in § 9 there is a brief discussion on the semi-group approach to the problem.

§ 2. Preliminary Theorems

The following known lemmas, needed in the succeeding sections, will be stated for completeness. The first concerns conditions under which a complex-valued function $h(v)$ of the real variable v admits a representation as a Fourier-Stieltjes integral.

Lemma 2.1. *Let $h(v)$ be integrable over every finite interval. A necessary and sufficient condition that $h(v)$ can be represented almost everywhere as*

$$h(v) = \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{\infty} e^{-ivx}\, dg(x), \tag{2.1}$$

where g is of bounded variation in $(-\infty, \infty)$, is that

$$\left\| \left(1/\sqrt{2\pi}\right) \int_{-R}^{R} \left(1 - \frac{|v|}{R}\right) e^{ivx} h(v)\, dv \right\|_{L_1} = O(1), \tag{2.2}$$

for all $0<R<\infty$. If $h(v)$ is continuous, the representation (2.1) *holds for all real v $(-\infty<v<\infty)$.*

This is a particular case of a general result due to H. Cramér [*19*] for which one may also refer to A. Gonzáles Dominguez [*26*] or J. L. B. Cooper [*17*, Theorem 5].

Needed also is the following fundamental result on the product of two Fourier-Stieltjes transforms, a proof of which may be found in D. V. Widder [*35*], p. 251—255.

Lemma 2.2. *If g_1 and g_2 are of bounded variation on $(-\infty, \infty)$ having Fourier-Stieltjes transforms $\check{g}_1(v)$ and $\check{g}_2(v)$ respectively, then*

$$\check{g}_1(v)\, \check{g}_2(v) = \check{\varphi}(v) \qquad (-\infty < v < \infty) \tag{2.3}$$

and

$$[\operatorname{Var} \varphi]_{-\infty}^{\infty} \leqq [\operatorname{Var} g_1]_{-\infty}^{\infty}\, [\operatorname{Var} g_2]_{-\infty}^{\infty}. \tag{2.4}$$

Here

$$\varphi(x) = \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{\infty} g_1(x-u)\, dg_2(u)$$

is the bilateral Stieltjes convolution of g_1 and g_2, and $[\mathrm{Var}\, g]_{-\infty}^{u}$ denotes the total variation of g over the interval $(-\infty, u)$.

In the case when $f \in L_p(-\infty, \infty)$, $1 < p \leqq 2$ a Fourier transform must be defined. Put

$$\hat{f}_\omega(v) = (1/\sqrt{2\pi}) \int_{-\omega}^{\omega} f(x)\, e^{-ivx}\, dx; \tag{2.5}$$

then as $\omega \uparrow \infty$, $\hat{f}_\omega(v)$ converges in $L_q(-\infty, \infty)$, $p+q=pq$, to a function $\hat{f}(v)$ (using the same notation as in L_1-space) called the *Fourier transform* of $f(x)$. This function is known to satisfy the inequality [*33*, p. 96]; [*37*, Vol. II, p. 254]

$$\left[(1/\sqrt{2\pi}) \int_{-\infty}^{\infty} |\hat{f}(v)|^q\, dv\right]^{1/q} \leqq \left[(1/\sqrt{2\pi}) \int_{-\infty}^{\infty} |f(x)|^p\, dx\right]^{1/p}. \tag{2.6}$$

This inequality is the extension to Fourier transforms of the classical inequality of Hausdorff-Young in the theory of Fourier series. In the case $p=q=2$, formula (2.6), with the inequality replaced by equality, is the classical PARSEVAL-PLANCHEREL formula.

The extension of Lemma 2.1 to L_p-functions is given by:

Lemma 2.3. *Let $h(v)$ be integrable over every finite interval. A necessary and sufficient condition that for almost all v*

$$h(v) = \underset{\omega\uparrow\infty}{\overset{(q)}{\mathrm{l.i.m.}}}\, (1/\sqrt{2\pi}) \int_{-\omega}^{\omega} g(x)\, e^{-ivx}\, dx,$$

where $g(x) \in L_p(-\infty, \infty)$, is that

$$\left\| (1/\sqrt{2\pi}) \int_{-R}^{R} \left(1 - \frac{|v|}{R}\right) h(v)\, e^{ivx}\, dv \right\|_{L_p} = O(1) \tag{2.7}$$

for all $0 < R < \infty$. (Here $\overset{(q)}{\mathrm{l.i.m.}}$ denotes the limit in the mean of order q.)

A proof of this lemma was kindly communicated to the author by J. L. B. COOPER, who will publish it.

Lemma 2.4. *If $f(x)$ and $f'(x)$ are absolutely continuous and*

$$g(x) = f''(x) \in L_p(-\infty, \infty), \qquad 1 < p \leqq 2,$$

then

$$\hat{g}(v) = (iv)^2 \hat{f}(v), \qquad \textit{almost all } v.$$

For this lemma, one may consult BOCHNER & CHANDRASEKHARAN [*5*, p. 124, p. 215] — it is the L_p-analogue of Theorem 3 (ii) of [*5*, p. 8].

The converse to this lemma is given by

Lemma 2.5. *If $f \in L_p(-\infty, \infty)$ and $g \in L_p(-\infty, \infty)$ for $1 < p \leqq 2$ and*

$$(iv)^2 \hat{f}(v) = \hat{g}(v), \qquad \textit{almost all } v,$$

then $f(x)$ and $f'(x)$ are absolutely continuous and $f''(x) \in L_p(-\infty, \infty)$.

For this result see [*5*, p. 128—129, p. 215].

Lemma 2.2 in L_p-space is given by:

Lemma 2.6. *Let* $g_1 \in L_1(-\infty, \infty)$ *having Fourier transform* $\hat{g}_1(v)$ *and* $g_2 \in L_p(-\infty, \infty)$, $1 < p \leqq 2$ *having Fourier transform* $\hat{g}_2(v) \in L_q(-\infty, \infty)$. *Then*

$$\hat{g}_1(v)\, \hat{g}_2(v) = \hat{\varphi}(v) \tag{2.8}$$

and

$$\|\varphi\|_{L_q} \leqq \|g_1\|_{L_1} \|g_2\|_{L_p}, \tag{2.9}$$

where

$$\varphi(x) = (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} g_1(x-u)\, g_2(u)\, du.$$

For a proof, one may consult Titchmarsh [*33*, Theorem 65, p. 90 in the case $p=2$ and Theorem 77, p. 106 for $1<p<2$].

§ 3. A General Class of Singular Integrals

Let f be a (non-periodic) function of class $L_1(-\infty, \infty)$, and let

$$J_\varrho(x) = (\varrho/\sqrt{2\pi}) \int_{-\infty}^{\infty} f(x+u)\, k(\varrho u)\, du \qquad (\varrho > 0) \tag{3.1}$$

be a general singular integral with kernel $k(u)$ having some or all of the properties:

i) $k(u)$ is a non-negative function of the real variable u, $-\infty < u < +\infty$ such that $k \in L_1(-\infty, \infty)$;

ii) $k(u)$ is continuous at $u=0$ and $k(u) = k(-u)$;

iii) $(1/\sqrt{2\pi}) \int_{-\infty}^{\infty} k(u)\, du = 1$;

and either

iv) $\mu_\gamma = (1/\sqrt{2\pi}) \int_{0}^{\infty} u^\gamma k(u)\, du < +\infty, \qquad \gamma > 0$

or the weaker condition

iv)* there exists a majorant k^* of k satisfying iv)

and either

v) $k(u)$ is monotone decreasing for $0 \leqq u < \infty$

or

v)* there exists a $k^* \in L_1(-\infty, \infty)$ such that $|k(u)| \leqq k^*(u)$ and k^* satisfies v).

Lemma 3.1. *Let* $f \in L_1(-\infty, \infty)$; *let* k *satisfy the conditions* i), ii) and iii). *Then* $J_\varrho(x)$ *as a function in* x *exists for almost all* x *and belongs to* $L_1(-\infty, \infty)$ *and*

$$\|J_\varrho\|_{L_1} \equiv (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} |J_\varrho(x)|\, dx \leqq \|f\|_{L_1}. \tag{3.2}$$

The proof is elementary; see also [*11*, §2].

Lemma 3.2. *Let* $f \in L_1(-\infty, \infty)$, *let* k *satisfy the conditions* i), ii) and iii). *Then*

$$\|J_\varrho(x) - f(x)\|_{L_1} \to 0, \qquad \varrho \uparrow \infty. \tag{3.3}$$

The proof is similar to [*33*, p. 35].

We now prove a direct theorem on the degree of approximation of $f(x)$ by the $J_\varrho(x)$ for large ϱ. It is the L_1-analogue of Theorem 2.3 of [11], the proof being entirely independent of Fourier transform theory. This theorem is to be compared with Theorem 5.1 of §5 below.

Theorem 3.1. *Assume that $f \in L_1(-\infty, \infty)$ and that k satisfies the conditions* i)—iii) *and* iv)* *and* v)* *(in particular* iv) *and* v)*). Let*

$$g(x, u) = f(x+u) + f(x-u) - 2f(x), \tag{3.4}$$

$$w(u) = \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{\infty} |g(x, u)|\, dx. \tag{3.5}$$

Then the condition

$$\int_0^h w(u)\, du = o(h^{1+\gamma}), \qquad h \to 0 \tag{3.6}$$

implies that

$$\|J_\varrho(x) - f(x)\|_{L_1} = 0(\varrho^{-\gamma}), \qquad \varrho \uparrow \infty. \tag{3.7}$$

Proof. Since the kernel k satisfies i) and ii), we have

$$\begin{aligned}\varrho^\gamma \int_{-\infty}^{\infty} |J_\varrho(x) - f(x)|\, dx &\leq \varrho^{1+\gamma} \int_{-\infty}^{\infty} dx \left\{\frac{1}{\sqrt{2\pi}} \int_0^\infty |g(x, u)|\, k^*(\varrho u)\, du\right\} \\ &= \varrho^{1+\gamma} \int_0^\infty k^*(\varrho u)\, du \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |g(x, u)|\, dx \\ &= \varrho^{1+\gamma} \int_0^\infty k^*(\varrho u)\, w(u)\, du = \varrho^{1+\gamma} \left(\int_0^\delta + \int_\delta^\infty\right) k^*(\varrho u)\, w(u)\, du \\ &= I_1 + I_2, \quad \text{say.}\end{aligned}$$

Defining

$$g^*(u) = \int_0^u w(v)\, dv,$$

we then integrate I_1 by parts, proceeding just as in the beginning of the proof of Theorem 2.1 of [11], and prove that to every $\varepsilon > 0$, there is a δ_0 such that

$$|I_1| < 3\,\varepsilon (1+\gamma) \sqrt{2\pi}\, \mu_\gamma, \qquad \delta \leqq \delta_0, \quad \text{all } \varrho > 0.$$

Then one fixes the δ and considers I_2.

Since $f \in L_1(-\infty, \infty)$, $w(u) \leqq 4\, \|f\|_{L_1} \equiv M$, it follows that

$$\begin{aligned}|I_2| &\leqq M \varrho^{1+\gamma} \int_\delta^\infty k^*(\varrho u)\, du \\ &\leqq \frac{M}{\delta^\gamma} \int_\delta^\infty (u\,\varrho)^\gamma\, k^*(\varrho u)\, \varrho\, du = \frac{M}{\delta^\gamma} \int_{\delta\varrho}^\infty v^\gamma\, k^*(v)\, dv \\ &< \frac{M}{\delta^\gamma} \sqrt{2\pi}\, \varepsilon, \qquad \varrho \geqq \varrho_0\end{aligned}$$

where we have used the fact that μ_γ is finite. Thus

$$\frac{\varrho^\gamma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |J_\varrho(x) - f(x)|\, dx \leqq \varepsilon \left\{3(1+\gamma)\,\mu_\gamma + \frac{M}{\delta^\gamma}\right\}$$

for $\varrho \geqq \varrho_0$ (δ being fixed). The proof is complete.

Theorem 3.2. *Let f and k satisfy the assumptions of Theorem 3.1 except that the condition (3.6) is replaced by*

$$\int_0^h w(u)\, du = O(h^{1+\gamma}), \qquad h \to 0. \tag{3.8}$$

Then

$$\|J_\varrho(x) - f(x)\|_{L_1} = O(\varrho^{-\gamma}), \qquad \varrho \uparrow \infty. \tag{3.9}$$

A slight modification of the proof of the previous theorem gives this result.

It may be remarked that the small-o as well as the large-O terms in (3.7) and (3.9), respectively, are independent of x.

It is important to note that Theorems 3.1 and 3.2 are of interest only in case $0 \leqq \gamma \leqq 2$. This is in some sense revealed by the following theorem, the method of proof of which is one particular instance of the Fourier-transform method explained in §1.

Theorem 3.3. *Let $f \in L_1(-\infty, \infty)$. If*

$$\lim_{t\to 0} \frac{1}{t^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |f(x+t) + f(x-t) - 2f(x)|\, dx = 0, \tag{3.10}$$

then $f(x) = 0$ for almost all x in $(-\infty, \infty)$.

Proof. Since $\mathfrak{F}[f(x+t)] = e^{ivt} \hat{f}(v)$, we have

$$\begin{aligned}\mathfrak{F}[f(x+t) + f(x-t) - 2f(x)] &= (e^{ivt} + e^{-ivt} - 2)\, \hat{f}(v)\\ &= (2\cos vt - 2)\, \hat{f}(v).\end{aligned}$$

Thus according to (3.10)

$$\begin{aligned}|2[\cos vt - 1]\, \hat{f}(v)| &\leqq \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{\infty} |[f(x+t) + f(x-t) - 2f(x)]\, e^{-ivx}|\, dx\\ &= o(t^2), \qquad t \to 0.\end{aligned}$$

It follows that

$$\lim_{t\to 0} \frac{1}{t^2} (\cos vt - 1)\, \hat{f}(v) = 0,$$

so that $(v^2/2!)\, \hat{f}(v) = 0$ for all v, and the uniqueness theorem for Fourier transforms implies that $f(x) = 0$ for almost all x in $(-\infty, \infty)$.

The proof reveals that the Fourier-transform method may sometimes be successfully applied to problems[2] which are not exactly of the type discussed in §1.

---

[2] See also BUTZER, On some theorems of Hardy, Littlewood and Titchmarsh (to appear).

§ 4. Inverse Approximation Theorems

First we prove a converse to Theorem 3.1.

Theorem 4.1. *Let $f \in L_1(-\infty, \infty)$; let k satisfy the conditions* i)—iii). *Suppose there exist constants $c > 0$ and, $0 < \gamma \leqq 2$ such that*

$$\lim_{\varrho \uparrow \infty} \frac{\varrho^\gamma}{|v|^\gamma} [1 - \hat{k}(v/\varrho)] = c > 0 \tag{4.1}$$

for every real v, $-\infty < v < \infty$. Then

$$\| J_\varrho(x) - f(x) \|_{L_1} = o(\varrho^{-\gamma}), \qquad (\varrho \uparrow \infty) \tag{4.2}$$

implies that $f(x) = 0$ for almost all x in $(-\infty, \infty)$.

Proof. We remark that in view of iii), $\hat{k}(0) = 1$ and $|\hat{k}(v)| \leqq 1$. Thus the constant c in (4.1) is positive.

We first note that in view of FUBINI's theorem it can readily be shown (see also [*11*], § 5) that

$$\mathfrak{F}[J_\varrho(x)] = \hat{f}(v)\, \hat{k}(v/\varrho) \qquad (-\infty < v < \infty).$$

This gives

$$\mathfrak{F}[f(x) - J_\varrho(x)] = \hat{f}(v)\, [1 - \hat{k}(v/\varrho)], \tag{4.3}$$

and according to (4.2)

$$|[1 - \hat{k}(v/\varrho)]\, \hat{f}(v)| \leqq (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} |[f(x) - J_\varrho(x)]\, e^{-ivx}|\, dx = o(\varrho^{-\gamma}), \qquad \varrho \uparrow \infty.$$

Applying (4.1) shows that

$$\lim_{\varrho \uparrow \infty} \varrho^\gamma [1 - \hat{k}(v/\varrho)]\, \hat{f}(v) = c\, |v|^\gamma \hat{f}(v) = 0$$

for all real v, and the uniqueness theorem for Fourier transforms gives $f(x) = 0$ for almost all x in $(-\infty, \infty)$. This establishes the theorem.

We note that the method of proof of the latter result is essentially that of Theorem 5.2 b) of [*11*]. The converse to Theorem 3.2 is given by

Theorem 4.2. *Let f and k satisfy the same assumptions as those of Theorem* 4.1, *including the condition* (4.1), *except that* (4.2) *is replaced by*

$$\| J_\varrho(x) - f(x) \|_{L_1} = O(\varrho^{-\gamma}) \qquad (\varrho \uparrow \infty). \tag{4.4}$$

Then there exists a function $g(x)$ of bounded variation on $(-\infty, \infty)$ such that $c|v|^\gamma \hat{f}(v) = \check{g}(v)$ for every real v.

Proof. We consider the partial integrals

$$s_w(x) = (1/\sqrt{2\pi}) \int_{-w}^{w} [1 - \hat{k}(v/\varrho)]\, \hat{f}(v)\, e^{ivx}\, dv, \tag{4.5}$$

which can be written as

$$s_w(x) = (1/\sqrt{2\pi}) \int_{-w}^{w} e^{ivx}\, dv\, (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} [f(u) - J_\varrho(u)]\, e^{-ivu}\, du.$$

Since the inner integral converges uniformly when $-w \leqq v \leqq +w$, where w is arbitrary, the interchange of the order of integration gives

$$
\begin{aligned}
s_w(x) &= (1/2\pi) \int_{-\infty}^{\infty} [f(u) - J_\varrho(u)]\, du \int_{-w}^{w} e^{iv(x-u)}\, dv \\
&= (1/\pi) \int_{-\infty}^{\infty} [f(u) - J_\varrho(u)] \frac{\sin w(x-u)}{(x-u)}\, du.
\end{aligned} \tag{4.6}
$$

We introduce the $(C, 1)$ means of the partial integrals $s_w(x)$, namely

$$
\sigma_R(x) = (1/R) \int_0^R s_w(x)\, dw = (1/\sqrt{2\pi}) \int_{-R}^{R} \left(1 - \frac{|v|}{R}\right) [1 - \hat{k}(v/\varrho)] \hat{f}(v)\, e^{ivx}\, dx. \tag{4.7}
$$

These Cesàro means $\sigma_R(x)$ may also be written, upon applying the second representation (4.6), in the form

$$
\sigma_R(x) = (1/R) \int_0^R \left\{\frac{1}{\pi} \int_{-\infty}^{\infty} [f(u) - J_\varrho(u)] \frac{\sin w(x-u)}{x-u}\, du\right\} dw,
$$

and inverting the order of integration in the repeated integral, we find that

$$
\sigma_R(x) = (1/\pi) \int_{-\infty}^{\infty} [f(u) - J_\varrho(u)] \frac{2 \sin^2 R(x-u)/2}{R(x-u)^2}\, du.
$$

In view of Lemma 2.1 and the hypothesis (4.3), it then follows that

$$
\|\sigma_R(x)\|_{L_1} \leqq \|f(x) - J_\varrho(x)\|_{L_1} = O(\varrho^{-\gamma}),
$$

yielding (by (4.7))

$$
\left\| (1/\sqrt{2\pi}) \int_{-R}^{R} \left(1 - \frac{|v|}{R}\right) [1 - \hat{k}(v/\varrho)] \hat{f}(v)\, e^{ivx}\, dv \right\|_{L_1} = O(\varrho^{-\gamma}) \tag{4.8}
$$

where the large-O term is independent of R and $\varrho > 0$.

On the other hand, by (4.1) it follows that there exists a constant $M > 0$ such that

$$
|\varrho^\gamma [1 - \hat{k}(v/\varrho)]| \leqq M |v|^\gamma \qquad (\varrho \geqq \varrho_0;\ |v| \leqq R),
$$

and thus

$$
\left| \left(1 - \frac{|v|}{R}\right) \varrho^\gamma [1 - \hat{k}(v/\varrho)] \hat{f}(v)\, e^{ivx} \right| \leqq 2M |v|^\gamma |\hat{f}(v)|
$$

for $\varrho \geqq \varrho_0$ and $|v| \leqq R$, every $R > 0$. Now the right-hand side of the latter inequality is integrable over every finite interval $|v| \leqq R$, and thus Lebesgue's dominated convergence theorem yields

$$
\lim_{\varrho \uparrow \infty} \int_{-R}^{R} \left(1 - \frac{|v|}{R}\right) \varrho^\gamma [1 - \hat{k}(v/\varrho)] \hat{f}(v)\, e^{ivx}\, dx = \int_{-R}^{R} \left(1 - \frac{|v|}{R}\right) c|v|^\gamma \hat{f}(v)\, e^{ivx}\, dx. \tag{4.9}
$$

By FATOU's lemma it follows that

$$\int_{-\infty}^{\infty}\left|(1/\sqrt{2\pi})\int_{-R}^{R}\left(1-\frac{|v|}{R}\right)c\,|v|^{\gamma}\hat{f}(v)\,e^{ivx}\,dv\right|dx$$

$$\leq \varliminf_{\varrho\uparrow\infty}\varrho^{\gamma}\int_{-\infty}^{\infty}(1/\sqrt{2\pi})\left|\int_{-R}^{R}\left(1-\frac{|v|}{R}\right)[1-\hat{k}(v/\varrho)]\,\hat{f}(v)\,e^{ivx}\,dv\right|dx$$

which is finite according to (4.8). Thus

$$\left\|(1/\sqrt{2\pi})\int_{-R}^{R}\left(1-\frac{|v|}{R}\right)c\,|v|^{\gamma}\hat{f}(v)\,e^{ivx}\,dv\right\|_{L_1}=O(1),$$

the large O being independent of R.

We finally apply Lemma 2.1 to the continuous function $h(v)=c|v|^{\gamma}\hat{f}(v)$ and deduce that $c|v|^{\gamma}\hat{f}(v)=\check{g}(v)$, all v. Thus the conclusion of the theorem follows.

We remark that an entirely different proof of the above theorem will be given in [*14*]. The above method has the advantage that it can readily be extended to L_p-functions for $1<p\leq 2$ (see § 6) and probably also to functions of several variables[3].

§ 5. The Fourier Method and Direct Theorems

This section is devoted to a direct approximation theorem which is essentially of the same type as Theorem 3.2. But here the condition (3.8) upon f will be replaced by one involving Fourier transforms of the function f and the conditions iv)* and v)* of the kernel k in §3 will be replaced by weaker conditions involving Fourier transforms. The method of proof of this theorem is also based upon the Fourier transform approach of this paper.

Theorem 5.1. *Let $f\in L_1(-\infty,\infty)$ and assume k satisfies* i)—iii) *of* § 3 *and instead of* (4.1) *the (apparently) stronger condition*

$$\frac{\varrho^{\gamma}[1-\hat{k}(v/\varrho)]}{c\,|v|^{\gamma}}=\check{\psi}(v/\varrho)\equiv(1/\sqrt{2\pi})\int_{-\infty}^{\infty}e^{-\frac{ivx}{\varrho}}\,d\psi(x),\tag{5.1}$$

where $c>0$, $0<\gamma\leq 2$ and $\psi(x)$ is of bounded variation on $(-\infty,\infty)$, and satisfying $\psi(-\infty)=0$ and $\psi(\infty)=\sqrt{2\pi}$. Then the condition

$$c\,|v|^{\gamma}\hat{f}(v)=\check{g}(v)\qquad(-\infty<v<\infty),$$

where $g(x)$ is of bounded variation on $(-\infty,\infty)$, implies that

$$\|J_{\varrho}(x)-f(x)\|_{L_1}=O(\varrho^{-\gamma}).$$

Proof. We have

$$\varrho^{\gamma}[1-\hat{k}(v/\varrho)]\,\hat{f}(v)=\check{g}(v)\,\check{\psi}(v/\varrho),$$

and by Lemma 2.2 it follows that

$$\check{g}(v)\,\check{\psi}(v/\varrho)=\check{\varphi}_{\varrho}(v)\qquad(-\infty<v<\infty),$$

[3] For a solution to the problem, see BUTZER, On DIRICHLET's problem for the half-space and the behavior of its solution on the boundary, Jour. of Mathematical Analysis and Applications, **1** (1960) (in press).

where

$$\check{\varphi}_\varrho(v) = \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{\infty} e^{-ivx}\, d\varphi_\varrho(x)$$

and

$$\varphi_\varrho(x) = \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{\infty} g(x-u)\, d\psi_\varrho(u)$$

anywhere we have put $\psi_\varrho(u) = \psi(\varrho u)$, $\varrho > 0$, $-\infty < u < \infty$. From (2.4) we obtain

$$\begin{aligned} [\operatorname{Var} \varphi_\varrho]_{-\infty}^{\infty} &\leqq \left(1/\sqrt{2\pi}\right) [\operatorname{Var} g]_{-\infty}^{\infty} [\operatorname{Var} \psi_\varrho]_{-\infty}^{\infty} \\ &\leqq \left(1/\sqrt{2\pi}\right) M_1 M_2 = M \end{aligned} \tag{5.2}$$

since both g and ψ are of bounded variation on $(-\infty, \infty)$. Here M_2 (and thus M) is independent of x, for if x ranges through $(-\infty, \infty)$, so does ϱx.

In view of (4.3), we deduce that

$$\varrho^\gamma \mathfrak{F}[f(x) - J_\varrho(x)] = \varrho^\gamma [1 - \hat{k}(v/\varrho)]\, \hat{f}(v) = \check{\varphi}_\varrho(v) = \mathfrak{F}\mathfrak{S}[\varphi_\varrho(x)].$$

Hence by the uniqueness theorem

$$\varphi_\varrho(x) = \varrho^\gamma \int_{-\infty}^{x} [f(u) - J_\varrho(u)]\, du.$$

The theorem on the variation of an indefinite Lebesgue integral (*e.g.* [*35*], p. 20) then gives

$$[\operatorname{Var} \varphi_\varrho]_{-\infty}^{\infty} = \varrho^\gamma \int_{-\infty}^{\infty} |f(x) - J_\varrho(x)|\, dx. \tag{5.3}$$

But according to (5.2), the total variation of φ_ϱ over the interval $(-\infty, \infty)$ is uniformly bounded, and thus the integral in (5.3) is also bounded by M. This establishes the theorem.

Remark. If the condition (5.1) is fulfilled, so is (4.1). Indeed,

$$\lim_{\varrho \uparrow \infty} \varrho^\gamma \frac{[1 - k(v/\varrho)]}{c|v|^\gamma} = \lim_{\varrho \uparrow \infty} \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{\infty} e^{-ivx/\varrho}\, d\psi(x) = \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{\infty} d\psi(x) = 1.$$

But whether the condition (5.1) is actually stronger than (4.1) remains unsolved. It may well be that both conditions are equivalent.

The above proof, which replaces an earlier incomplete argument of the author, was established in conjunction with Heinz König, who introduced the condition (5.1). The author is grateful for these kind suggestions. The metric (5.3) was introduced into analysis by S. Bochner [*4*].

It would be of interest to express the condition (4.1) directly in terms of properties of the kernel k itself. This question is now being investigated by H. König [*29*], who has just proven the following result:

Lemma 5.1. *Let k satisfy* i)—iii) *and* $0 < \gamma < 2$. *Then the condition* (4.1) *is equivalent to*

$$\lim_{t \uparrow \infty} t^\gamma \int_t^{\infty} k(u)\, du = \sqrt{\frac{2}{\pi}}\, c\, \Gamma(\gamma) \sin \frac{\pi\gamma}{2}.$$

If in addition k satisfies v), *then* (4.1) *is equivalent to*

$$\lim_{t\uparrow\infty} t^{\gamma+1} k(t) = \sqrt{\frac{2}{\pi}}\, c\, \Gamma(\gamma+1) \sin\frac{\pi\gamma}{2}.$$

A comparison of the hypotheses of Theorems 3.2 and 5.1 is also of interest. We first consider the case $\gamma=1$ and compare the conditions iv) of §3 with (4.1) and (5.1).

If

$$2\mu_1 = (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} |x|\, k(x)\, dx < \infty, \tag{5.4}$$

then the existence of

$$\lim_{\varrho\uparrow\infty} \varrho \frac{[1-\hat{k}(v/\varrho)]}{|v|} = \lim_{\varrho\uparrow\infty} \frac{\varrho}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[\frac{1-e^{-ivx/\varrho}}{|v x|}\right] |x|\, k(x)\, dx = 0 \tag{5.5}$$

follows. But the converse does not hold in general. Indeed, consider FEJÉR'S kernel of § 7; the limit in (5.5) exists as well as (5.1), but the first moment in (5 4) does not exist. Thus the assumptions of Theorem 3.2 seem to be stronger than those of Theorem 5.1 in the case $\gamma=1$.

But in the case $\gamma=2$, this is not quite so. If the second moment

$$2\mu_2 = (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} x^2 k(x)\, dx$$

is finite, then

$$\lim_{\varrho\uparrow\infty} \varrho^2 \frac{[\hat{k}(v/\varrho)-2+\hat{k}(-v/\varrho)]}{v^2} = -(1/\sqrt{2\pi}) \int_{-\infty}^{\infty} x^2 k(x)\, dx. \tag{5.6}$$

Conversely, if the limit in (5.6) exists, so does the second moment (compare [6], p. 71). Thus the essential difference between Theorems 3.2 and 5.1 is that the assumptions (3.8) and v) (or v)*) of § 3 must be compared with the condition $c\, v^2 \hat{f}(v) = \check{g}(v)$, for all v, of Theorem 5.1.

Theorem 5.2. *Let*

$$J_\varrho(x) = \frac{\varrho}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x+u)\, k(\varrho u)\, du$$

be a given singular integral corresponding to a function $f \in L_1(-\infty,\infty)$, where the kernel k satisfies the conditions i)—iii) *of* § 3 *and also the condition* (5.1). *Then the integral $J_\varrho(x)$ is said to be saturated with order $O(\varrho^{-\gamma})$, and the saturation class is the class of functions f for which*

$$c\,|v|^\gamma \hat{f}(v) = \check{g}(v) \qquad (-\infty < v < \infty) \tag{5.4}$$

where g is of bounded variation on $(-\infty,\infty)$, and $c>0$, $0<\gamma\leqq 2$.

In other words, a necessary and sufficient condition that the difference $D_\varrho(J; f)$ is exactly of order $O(\varrho^{-\gamma})$ is that the condition (5.4) be satisfied.

The proof follows immediately from Theorems 4.1, 4.2, 5.1 and the definition of a saturation class given in §1.

§ 6. Approximation in the Space $L_p(-\infty, \infty)$, $1<p\leqq 2$

In this section we proceed to extend the foregoing results from L_1 to L_p-space. For this purpose we have already introduced the Fourier transform in L_p in §2.

Lemma 6.1. *Let $f\in L_p(-\infty, \infty)$, $p\geqq 1$; let k satisfy the conditions* i), ii) *and* iii).

a) *Then the integral $J_\varrho(x)$ exists for almost all x and belongs to $L_p(-\infty, \infty)$, and*

$$\|J_\varrho(x)\|_{L_p}\equiv\left((1/\sqrt{2\pi})\int_{-\infty}^{\infty}|J_\varrho(x)|^p\,dx\right)^{1/p}\leqq\|f\|_{L_p}.$$

b) *Also*

$$\|J_\varrho(x)-f(x)\|_{L_p}\to 0\qquad(\varrho\uparrow\infty).$$

For the proof one may consult [*5*, p. 98—101].

The extension to L_p-functions of Theorems 4.1 and 4.2 is the following.

Theorem 6.1. *Let $f\in L_p(-\infty, \infty)$, $1<p\leqq 2$, let k satisfy the conditions* i)—iii), *and suppose there exist constants $c>0$, $0<\gamma\leqq 2$ satisfying* (4.1) *for every real v.*

a) *If*

(6.1) $$\|J_\varrho(x)-f(x)\|_{L_p}=o(\varrho^{-\gamma})\qquad(\varrho\uparrow\infty),$$

then $f(x)=0$ a.e. in $(-\infty, \infty)$.

b) *If*

(6.2) $$\|J_\varrho(x)-f(x)\|_{L_p}=O(\varrho^{-\gamma}),$$

then $c|v|^\gamma\hat{f}(v)$ is the Fourier transform of a function g in $L_p(-\infty, \infty)$.

Proof. a) It can readily be shown (compare [*37*, Vol. II, p. 253]) that

$$\mathfrak{F}[J_\varrho(x)]=f(v)\,\hat{k}(v/\varrho),$$

and thus in view of the Hausdorff-Young inequality (see § 2)

$$\left[(1/\sqrt{2\pi})\int_{-\infty}^{\infty}|\hat{f}(v)\,[1-\hat{k}(v/\varrho)]|^q\,dv\right]^{1/q}\leqq\left[(1/\sqrt{2\pi})\int_{-\infty}^{\infty}|f(x)-J_\varrho(x)|^p\,dx\right]^{1/p}.$$

The hypothesis (6.1) gives

$$\lim_{\varrho\uparrow\infty}\varrho^\gamma\left[(1/\sqrt{2\pi})\int_{-\infty}^{\infty}|\hat{f}(v)\,[1-\hat{k}(v/\varrho)]|^q\,dv\right]^{1/q}=0.$$

From (4.1) and FATOU's lemma it follows that

$$\left[(1/\sqrt{2\pi})\int_{-\infty}^{\infty}c^q\,|v^\gamma\hat{f}(v)|^q\,dv\right]^{1/q}=0,$$

and thus $c|v|^\gamma\hat{f}(v)=0$ for almost all points in $(-\infty, \infty)$. Only application of the uniqueness theorem is needed to show that $f(x)=0$ a.e. in $(-\infty, \infty)$.

b) We consider the partial integrals

$$s_w(x)=(1/\sqrt{2\pi})\int_{-w}^{w}\hat{f}(v)\,[1-\hat{k}(v/\varrho)]\,e^{ivx}\,dv\qquad(0<w<\infty).$$

The transform of $\hat{l}(v)=e^{ivx}(|v|<w)$, $0(|v|>w)$, is

$$l(u)=(1/\sqrt{2\pi})\int_{-w}^{w}e^{ivx}\,e^{ivu}\,dv=\sqrt{\frac{2}{\pi}}\,\frac{\sin w(x+u)}{(x+u)}.$$

Hence, by PARSEVAL's formula [33, p. 70 for $p=2$, p. 105 for $1<p<2$]

$$(1/\sqrt{2\pi})\int_{-w}^{w}\hat{f}(v)\,[1-\hat{k}(v/\varrho)]\,e^{ivx}\,dv=(1/\pi)\int_{-\infty}^{\infty}[f(u)-J_\varrho(u)]\,\frac{\sin w(x-u)}{x-u}\,du.$$

It follows that

$$s_w(x)=(1/\pi)\int_{-\infty}^{\infty}[f(u)-J_\varrho(u)]\,\frac{\sin w(x-u)}{x-u}\,du.$$

As in the proof of Theorem 4.2, we now consider

$$\sigma_R(x)=(1/R)\int_0^R s_w(x)\,dw=(1/\sqrt{2\pi})\int_{-R}^{R}\left(1-\frac{|v|}{R}\right)\hat{f}(v)\,[1-\hat{k}(v/\varrho)]\,e^{ivx}\,dv.$$

The transform of

$$\hat{m}(v)=\begin{cases}\left(1-\dfrac{|v|}{R}\right)e^{ivx} & (|v|<R)\\ 0 & (|v|>R),\end{cases}$$

is

$$m(u)=(1/\sqrt{2\pi})\int_{-R}^{R}\left(1-\frac{|v|}{R}\right)e^{ivx}\,e^{ivu}\,dv=\sqrt{\frac{2}{\pi}}\,\frac{\sin^2 R(x+u)/2}{R(x+u)^2}.$$

Thus again by PARSEVAL's formula

$$\begin{aligned}(1/\sqrt{2\pi})&\int_{-R}^{R}\hat{f}(v)\,[1-\hat{k}(v/\varrho)]\left(1-\frac{|v|}{R}\right)e^{ivx}\,dv\\ &=(1/\pi)\int_{-\infty}^{\infty}[f(u)-J_\varrho(x)]\,\frac{\sin^2 R(x-u)/2}{R(x-u)^2}\,du.\end{aligned}$$

In view of Lemma 6.1 and the hypothesis (6.2) we have

$$\|\sigma_R(x)\|_{L_p}\leqq\|f(x)-J_\varrho(x)\|_{L_p}=O(\varrho^{-\gamma}),$$

giving

$$\left\|(1/\sqrt{2\pi})\int_{-R}^{R}\left(1-\frac{|v|}{R}\right)\hat{f}(v)\,[1-\hat{k}(v/\varrho)]\,e^{ivx}\,dv\right\|_{L_p}=O(\varrho^{-\gamma}),\tag{6.3}$$

the large-O term being independent of R and $\varrho>0$.

We proceed now as in the proof of Theorem 4.2 obtaining (4.9). By FATOU's lemma it then follows that

$$\begin{aligned}&\left\|(1/\sqrt{2\pi})\int_{-R}^{R}\left(1-\frac{|v|}{R}\right)c\,|v|^\gamma\hat{f}(v)\,e^{ivx}\,dv\right\|_{L_p}\\ &\qquad\leqq\varliminf_{\varrho\uparrow\infty}\varrho^\gamma\left\|(1/\sqrt{2\pi})\int_{-R}^{R}\left(1-\frac{|v|}{R}\right)\hat{f}(v)\,[1-\hat{k}(v/\varrho)]\,e^{ivx}\,dv\right\|_{L_p}<+\infty.\end{aligned}$$

Finally, applying Lemma 2.3 to the function $h(v)=c|v|^\gamma \hat{f}(v)$, we deduce that

(6.4) $$c|v|^\gamma \hat{f}(v) = \underset{\omega\uparrow\infty}{\overset{(q)}{\text{l.i.m.}}}\ (1/\sqrt{2\pi}) \int_{-\omega}^{\omega} g(x)\, e^{-ivx}\, dx$$

for almost all v, where $g\in L_p(-\infty,\infty)$. This completes the proof.

Theorem 6.2. *Let* $f\in L_p(-\infty,\infty)$, $1<p\leqq 2$, *and assume* k *satisfies* i)—iii) *of* §3 *and*

(6.5) $$\frac{\varrho^\gamma[1-\hat{k}(v/\varrho)]}{c|v|^\gamma} = \check{\psi}(v/\varrho) \equiv (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} e^{-ivx/\varrho}\, d\psi(x),$$

where $c>0$, $0<\gamma\leqq 2$,

$$\psi(x) = \int_{-\infty}^{x} \alpha(u)\, du \qquad (-\infty<x<\infty)$$

and

$$\alpha(x)\in L_1(-\infty,\infty).$$

Then the condition

(6.6) $$c|v|^\gamma \hat{f}(v) = \hat{g}(v)$$

where $g(x)\in L_p(-\infty,\infty)$, *implies that*

$$\|J_\varrho(x)-f(x)\|_{L_p} = O(\varrho^{-\gamma}).$$

Proof. We have

$$\varrho^\gamma[1-\hat{k}(v/\varrho)]\hat{f}(v) = \check{\psi}(v/\varrho)\,\hat{g}(v) = (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} \varrho\,\alpha(\varrho u)\, e^{-ivu}\, du\cdot \hat{g}(v).$$

By Lemma 2.6 it follows that the product of the two transforms equals

$$\hat{\varphi}_\varrho(v) = \mathfrak{F}[\varphi_\varrho(x)], \qquad \varphi_\varrho(x) = \frac{\varrho}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(x-u)\,\alpha(\varrho u)\, du.$$

Also

(6.7) $$\|\varphi_\varrho\|_{L_p} \leqq \|g\|_{L_p}\, \varrho\, \|\alpha_\varrho\|_{L_1},$$

where we have put $\alpha_\varrho(u)=\alpha(\varrho u)$, $\varrho>0$.

Now

$$\psi(\varrho x) = \varrho \int_{-\infty}^{x} \alpha(\varrho u)\, du,$$

and thus

$$(1/\sqrt{2\pi})\, [\operatorname{Var}\psi_\varrho]_{-\infty}^{\infty} = \varrho\, \|\alpha_\varrho\|_{L_1},$$

where we have put $\psi_\varrho(u)=\psi(\varrho u)$, $\varrho>0$. According to (6.7) it then follows that

$$\|\varphi_\varrho\|_{L_p} \leqq (1/\sqrt{2\pi})\, \|g\|_{L_p}\, [\operatorname{Var}\psi_\varrho]_{-\infty}^{\infty} \leqq M_1 M_2 = M,$$

since $g\in L_p(-\infty,\infty)$ and ψ is of bounded variation. Also M_2 is independent of ϱ since ϱx varies through $(-\infty,\infty)$ if x does. Now

$$\varrho^\gamma\, \mathfrak{F}[f(x)-J_\varrho(x)] = \varrho^\gamma[1-\hat{k}(v/\varrho)]\hat{f}(v) = \hat{\varphi}_\varrho(v) = \mathfrak{F}[\varphi_\varrho(x)].$$

By the uniqueness theorem for L_p-transforms we have

$$\varrho^\gamma [f(x) - J_\varrho(x)] = \varphi_\varrho(x).$$

Since $\|\varphi_\varrho\|_{L_p} \leqq M$, we find that

$$\varrho^\gamma \|f(x) - J_\varrho(x)\|_{L_p} \leqq M,$$

where M is independent of ϱ, establishing the theorem.

As an immediate conclusion of these theorems we have

Theorem 6.3. *Let $f \in L_p(-\infty, \infty)$, $1 < p \leqq 2$, and assume k satisfies the conditions* i)—iii) *of* §3 *as well as the condition* (6.5) *with $\psi(x)$ and $\alpha(x)$ as defined there.*

a) *If $\|f(x) - J_\varrho(x)\|_{L_p} = o(\varrho^{-\gamma})$, $\varrho \uparrow \infty$, then $f = 0$* a.e.

b) *The following statements are equivalent for $J_\varrho(x)$:*

i) $\|f(x) - J_\varrho(x)\|_{L_p} = O(\varrho^{-\gamma}) \qquad (\varrho \uparrow \infty)$;

ii) *there is a g in $L_p(-\infty, \infty)$ so that $\hat{g}(v) = c|v|^\gamma \hat{f}(v)$ for almost all v.*

In § 9 this theorem will be compared with a fundamental theorem on semi-groups in the case $J_\varrho(x)$ is a semi-group operator.

§ 7. Applications to Various Singular Integrals

Consider *Fejér's integral:*

$$\sigma_n(x) = \frac{1}{2\pi(n+1)} \int_{-\infty}^{\infty} \left[\frac{\sin(n+1)\,u/2}{u/2}\right]^2 f(x+u)\,du \tag{7.1}$$

for $f \in L_1(-\infty, \infty)$, and let $D_n(\sigma; f) = \|\sigma_n(x) - f(x)\|_{L_1}$.

As a first application of the general theorems established above, we have the following

Theorem 7.1. *Let $f \in L_1(-\infty, \infty)$.*

a) *If $D_n(\sigma; f) = o(1/n)$, $n \uparrow \infty$, then f is zero almost everywhere in $(-\infty, \infty)$.*

b) *If $D_n(\sigma; f) = O(1/n)$, then there exists a function g of bounded variation on $(-\infty, \infty)$ such that*

$$|v|\,\hat{f}(v) = \check{g}(v) \qquad (-\infty < v < \infty). \tag{7.2}$$

c) *If the condition* (7.2) *is satisfied, g being a function of bounded variation on $(-\infty, \infty)$, then $D_n(\sigma; f) = O(1/n)$.*

In other words, Fejér's integral is saturated with order $O(1/n)$, and its saturation class is the class of functions f for which $|v|\,\hat{f}(v)$ is the Fourier-Stieltjes integral of a function g of bounded variation on $(-\infty, \infty)$.

Proof. The kernel $\hat{k}$ with its Fourier transform k, in case of Fejér's integral, is given by

$$k(u) = \frac{1}{\sqrt{2\pi}} \left(\frac{\sin u/2}{u/2}\right)^2, \qquad \hat{k}(v) = \begin{cases} 1 - |v| & (|v| \leqq 1) \\ 0 & (|v| \geqq 1). \end{cases}$$

Here $\varrho=(n+1)$, $\gamma=1$, and the condition (4.1) takes on the form

$$\lim_{n\uparrow\infty}\frac{(n+1)}{|v|}\left[1-\hat{k}\left(\frac{v}{n+1}\right)\right]=1 \qquad (-\infty<v<\infty) \tag{7.3}$$

with constant c equal to one. Upon applying Theorems 4.1 and 4.2, parts a) and b) then follow immediately.

To prove part c) one must show (5.1) is satisfied. Indeed, according to Fourier transform tables (*e.g.* ERDÉLYI *et al.* [22] p. 18 and p. 41)

$$\sqrt{\frac{2}{\pi}}\int_0^\infty\left\{\sqrt{\frac{2}{\pi}}\left(\frac{\sin x}{x}-Ci(x)\right)\right\}\cos\left(x\frac{v}{n+1}\right)dx=\begin{cases}1 & \text{for } 0<\dfrac{v}{n+1}<1\\[2mm] \dfrac{n+1}{v} & \text{for } \dfrac{v}{n+1}>1,\end{cases}$$

where

$$Ci(x)=-\int_x^\infty\frac{\cos u}{u}\,du$$

is the cosine integral. This gives

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty\left\{\sqrt{\frac{2}{\pi}}\left(\frac{\sin x}{x}-Ci(x)\right)\right\}e^{-ix\frac{v}{n+1}}\,dx=\begin{cases}1 & \text{for } \dfrac{|v|}{n+1}\leqq 1\\[2mm] \dfrac{n+1}{|v|} & \text{for } \dfrac{|v|}{n+1}\geqq 1.\end{cases}$$

Hence the quotient

$$\frac{n+1}{|v|}\left[1-\hat{k}\left(\frac{v}{n+1}\right)\right]=\begin{cases}1 & \text{for } \dfrac{|v|}{n+1}\leqq 1\\[2mm] \dfrac{n+1}{|v|} & \text{for } \dfrac{|v|}{n+1}\geqq 1,\end{cases}$$

may be expressed as a Fourier-Stieltjes integral with

$$\psi(x)=\sqrt{\frac{2}{\pi}}\int_{-\infty}^x\left(\frac{\sin u}{u}-Ci(u)\right)du,$$

where $\psi(-\infty)=0$ and $\psi(\infty)=\sqrt{2\pi}$, since

$$\int_{-\infty}^\infty\frac{\sin u}{u}\,du=\pi, \qquad \int_{-\infty}^\infty Ci(u)\,du=0,$$

so that the theorem is established.

It would naturally be interesting to express the condition (7.2), given in terms of Fourier transforms, directly in terms of the original function space. This leads to the following result, which I have not been able to locate in the literature:

Conjecture A: a) *If* $f\in L_1(-\infty,\infty)$ *and if* (7.2) *is satisfied for every real* v *where* g *is of bounded variation on* $(-\infty,\infty)$, *then the conjugate function* $\tilde{f}(x)$ *is of bounded variation on* $(-\infty,\infty)$.

b) *If* $f\in L_1(-\infty,\infty)$ *and* $g(x)=\tilde{f}(x)$ *is of bounded variation on* $(-\infty,\infty)$, *then*

$$\check{g}(v)=|v|\,\hat{f}(v) \qquad (-\infty<v<\infty).$$

Here

$$\tilde{f}(x) = -\lim_{\varepsilon \downarrow 0} \frac{1}{\pi} \int_{\varepsilon}^{\infty} \frac{f(x+u) - f(x-u)}{u} \, du.$$

Supposing that this conjecture can be verified, it would follow that the saturation class of Fejér's integral is the class of functions f for which $\tilde{f}(x)$ is of bounded variation on $(-\infty, \infty)$.

Let us now proceed to the case $f \in L_p(-\infty, \infty)$, $1 < p \leqq 2$. An application of Theorems 6.1, 6.2 gives

Theorem 7.2. *Let f be in $L_p(-\infty, \infty)$, $1 < p \leqq 2$. Then the following statements are equivalnt for* Fejér's *integral:*

1. $\|\sigma_n(x) - f(x)\|_{L_p} = O(1/n) \qquad (n \uparrow \infty)$;
2. *there is a g in $L_p(-\infty, \infty)$ so that $\hat{g}(v) = |v| \, \hat{f}(v)$ for almost all v.*

Furthermore, if $\|\sigma_n(x) - f(x)\|_{L_p} = o(1/n)$, then f must be zero almost everywhere in $(-\infty, \infty)$.

We may remark that results analogous to those of Theorems 7.1 and 7.2 have occured in the literature (*e.g.* Alexits [*1*]) only in the case when f is a periodic function of period 2π, $f \in L_p(-\pi, \pi)$, $p \geqq 1$. For the proof of part a) of Theorem 7.1 see also [*11*, § 6].

It may be mentioned that the semi-group methods of [*8*, *9*] cannot be applied directly to Fejér's integral, since it is not a semi-group operator.

The *Poisson-Cauchy singular integral* is defined by

$$p(x, y) = \frac{y}{\pi} \int_{-\infty}^{\infty} \frac{f(x+u)}{y^2 + u^2} \, du,$$

with

$$k(u) = \sqrt{\frac{2}{\pi}} \, \frac{1}{1+u^2}, \qquad \hat{k}(v) = e^{-|v|}.$$

Here $\varrho = 1/y$, $\gamma = 1$, and (4.1) takes on the form

$$\lim_{y \downarrow 0} \frac{1}{y|v|} \left[1 - e^{-y|v|}\right] = 1 \qquad (-\infty < v < \infty).$$

Regarding the condition (5.1), we note that

$$\frac{1}{y|v|} \left[1 - e^{-y|v|}\right]$$

may be expressed as a Fourier-Stieltjes integral with

$$\psi(x) = \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{x} \log\left(\frac{u^2+1}{u^2}\right) du,$$

where $\psi(-\infty) = 0$ and $\psi(\infty) = \sqrt{2\pi}$. This follows from the fact that ([*22*], p. 18)

$$\sqrt{\frac{2}{\pi}} \int_{0}^{\infty} \left\{ \frac{1}{\sqrt{2\pi}} \log \frac{x^2+1}{x^2} \right\} \cos(v y x) \, dx = \frac{1 - e^{-y|v|}}{y|v|} \qquad (v > 0).$$

From Theorems 4.1, 4.2 and 5.1 then follows readily

Theorem 7.3. *Let* $f \in L_1(-\infty, \infty)$, $D_y(p; f) = \|p(x, y) - f(x)\|_{L_1}$. *Then the following statements are equivalent for the Poisson-Cauchy integral:*

1. $D_y(p, f) = O(y) \qquad (y \uparrow 0)$;

2. *There is a function* g *of bounded variation in* $(-\infty, \infty)$ *so that* $\check{g}(v) = |v| \hat{f}(v)$, *all* v.

Furthermore, if $D_y(p; f) = o(y)$, $y \downarrow 0$, *then* f *is zero a.e. in* $(-\infty, \infty)$.

In the case of L_p-functions we have

Theorem 7.4. *If* $f \in L_p(-\infty, \infty)$, $1 < p \leqq 2$, *then the Poisson-Cauchy integral is saturated with order* $O(y)$, $y \downarrow 0$, *and its saturation class is the class of functions* f *for which* $|v| \hat{f}(v)$ *is the Fourier transform of a function* $g \in L_p(-\infty, \infty)$.

The *Gauss-Weierstrass singular integral*, given by

$$W(t; x) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} f(x+u)\, e^{-u^2/4t}\, du,$$

has its kernel k with corresponding transform given by

$$k(u) = \sqrt{2}\, e^{-u^2}, \qquad \hat{k}(v) = e^{-v^2/4}.$$

Here $\varrho = (2\sqrt{t})^{-1}$, $\gamma = 2$, and (4.1) is of the form

$$\lim_{t \downarrow 0} \frac{1}{t v^2} [1 - e^{-t v^2}] = 1 \qquad (-\infty < v < \infty).$$

Concerning (5.1), we show that

$$\frac{1}{t v^2} [1 - e^{-t v^2}]$$

can be expressed as a Fourier-Stieltjes transform with

$$\psi(x) = 2\sqrt{2} \int_{-\infty}^{x} (e^{-u^2} - \sqrt{\pi}\, u \operatorname{Erfc} u)\, du,$$

where

$$\operatorname{Erfc} x = (2/\sqrt{\pi}) \int_{x}^{\infty} e^{-u^2}\, du,$$

is the complementary error function. This follows from the known formula ([22], p. 15 and p. 40)

$$\sqrt{\frac{2}{\pi}} \int_0^{\infty} \{2\sqrt{2}(e^{-x^2} - \sqrt{\pi}\, x \operatorname{Erfc} x)\} \cos(2\sqrt{t}\, v x)\, dx = \frac{1}{t v^2} [1 - e^{-t v^2}] \quad (v > 0).$$

We also note that $\psi(-\infty) = 0$, $\psi(\infty) = \sqrt{2\pi}$, the latter following from

$$\int_{-\infty}^{\infty} u \operatorname{Erfc} u\, du = 0.$$

Theorem 7.5. *Let $f \in L_1(-\infty, \infty)$. Then the following statements are equivalent for the Gauss-Weierstrass integral:*

1. $\|W(t; x) - f(x)\|_{L_1} = O(t) \qquad (t \downarrow 0)$;

2. *there is a g of bounded variation on $(-\infty, \infty)$ so that $\check{g}(v) = v^2 \hat{f}(v)$, for all v. Furthermore, if $\|W(t; x) - f(x)\|_{L_1} = o(t)$, then f is zero a.e.*

In the case of L_p-functions, we may even characterize the saturation class in terms of the original function-space; thus we have

Theorem 7.6. *If $f \in L_p(-\infty, \infty)$, $1 < p \leqq 2$, then the Gauss-Weierstrass singular integral is saturated with order $O(t)$, $t \downarrow 0$, and its saturation class is the class of functions f for which f and f' are absolutely continuous and the second derivative $f''(x)$ is in $L_p(-\infty, \infty)$.*

Proof. From Theorems 6.1, 6.2 and 6.3 one obtains the saturation class expressed in terms of Fourier transforms. Then one makes use of Lemmas 2.4 and 2.5 to obtain the characterization of the saturation class in terms of the original functions.

Remark. The Poisson-Cauchy integrals as well as the Gauss-Weierstrass integrals are semi-group operators. Thus semi-group theory may be applied to give entirely different proofs of Theorems 7.3—7.6. In the case of the reflexive Banach space $L_p(-\infty, \infty)$, $1 < p \leqq 2$, these have been given by the author [*8*, p. 421—423] and for L_1-space by K. DE LEEUW [*30*, § 5, 6]. In the case of non-reflexive Banach spaces, K. DE LEEUW [*30*] has presented an elegant generalization of the semi-group methods of [*8, 9*]. But the Fourier transform method seems to have many advantages, especially in L_1-space. It is shorter and more direct.

§ 8. Further Applications

The object of this section is to present two further singular integrals which are not semi-group operators, so that semi-group methods cannot be applied directly to them.

Consider PICARD's *singular integral* (see also [*11*, § 8])

$$P(x; t) = \frac{1}{2t} \int_{-\infty}^{\infty} f(x+u)\, e^{-|u|/t}\, du,$$

with

$$k(u) = \sqrt{\frac{\pi}{2}}\, e^{-|u|}, \qquad \hat{k}(v) = \frac{1}{1+v^2}.$$

Theorem 8.1. *If $f \in L_1(-\infty, \infty)$, then* PICARD's *integral is saturated with order $O(t^2)$, $t \downarrow 0$ and its saturation class is the class of functions f for which $v^2 \hat{f}(v)$ is the Fourier-Stieltjes transform of a function g of bounded variation on $(-\infty, \infty)$.*

Proof. For this singular integral $\varrho = 1/t$, $\gamma = 2$, and (4.1) is of the form

$$\lim_{t \downarrow 0} \frac{1}{v^2 t^2} \left[1 - \frac{1}{1+v^2 t^2}\right] = 1 \qquad (-\infty < v < \infty).$$

We also note that

$$\frac{1}{v^2 t^2} \left[1 - \frac{1}{1+v^2 t^2}\right] = \frac{1}{1+(v t)^2}$$

may be expressed as a Fourier-Stieltjes integral with

$$\psi(x) = \sqrt{\frac{\pi}{2}} \int_{-\infty}^{x} e^{-|u|}\, du,$$

where $\psi(-\infty)=0$ and $\psi(\infty)=\sqrt{2\pi}$. By Theorem 5.2 the proof is complete.

Conjecture B: a) *If $f \in L_1(-\infty, \infty)$, g is of bounded variation in $(-\infty, \infty)$ and*

$$(i v)^2 \hat{f}(v) = \check{g}(v), \qquad \textit{for all } v,$$

then $f'(x)$ is of bounded variation on $(-\infty, \infty)$.

b) *If $f \in L_1(-\infty, \infty)$ and $g(x)=f'(x)$ is of bounded variation on $(-\infty, \infty)$, then*

$$\check{g}(v) = (i v)^2 \hat{f}(v), \qquad \textit{for all } v.$$

Now if this conjecture holds, *the saturation class of* Picard's *integral is the class of functions f for which $f'(x)$ is of bounded variation on $(-\infty, \infty)$.*

For L_p-functions we have

Theorem 8.2. *Let $f \in L_p(-\infty, \infty)$, $1<p\leqq 2$. For* Picard's *integral the following statements are equivalent:*

1. $\|P(x;t) - f(x)\|_{L_p} = O(t^2)$, $(t \downarrow 0)$;
2. *f and f' are absolutely continuous and f'' is in $L_p(-\infty, \infty)$.*

Furthermore, if $\|P(x;t) - f(x)\|_{L_p} = o(t^2)$, then f is zero a.e.

For the *Jackson-de la Vallée Poussin integral* corresponding to the function f

$$J_{2n}(x) = \frac{3}{4\pi(n+1)^3} \int_{-\infty}^{\infty} \left[\frac{\sin(n+1)\,u/2}{u/2}\right]^2 f(x+u)\, du,$$

we have

Theorem 8.3. *Assume $f \in L_1(-\infty, \infty)$.* Jackson's *integral is saturated with order $O(1/n^2)$, $n \uparrow \infty$, and its saturation class consists of those functions f for which $(3/2)\, v^2 \hat{f}(v)$ is the Fourier-Stieltjes transform of a function g of bounded variation on $(-\infty, \infty)$.*

Proof. Here $\varrho=(n+1)$, $\gamma=2$

$$k(u) = \frac{3}{2\sqrt{2\pi}} \left(\frac{\sin u/2}{u/2}\right)^4, \qquad \hat{k}(v) = \begin{cases} 1 - \frac{3}{2} v^2 + \frac{3}{4}|v|^3 & (|v| \leqq 1) \\ \frac{1}{4}(2 - |v|^3) & (1 \leqq |v| \leqq 2) \\ 0 & (|v| \geqq 2), \end{cases}$$

and

$$\lim_{n\uparrow\infty} \frac{(n+1)^2}{\left(\frac{3}{2}\right) v^2} \left[1 - \hat{k}\left(\frac{v}{n+1}\right)\right] = 1 \qquad \text{for all } v.$$

Also

$$\frac{(n+1)^2}{\left(\frac{3}{2}\right) v^2}\left[1 - \hat{k}\left(\frac{v}{n+1}\right)\right] = \begin{cases} 1 - \frac{1}{2}\left|\frac{v}{n+1}\right|, & \text{for} \quad \left|\frac{v}{n+1}\right| \leqq 1 \\ \frac{2}{3}\left(\frac{n+1}{v}\right)^2 \left\{1 - \frac{1}{4}\left(2 - \left|\frac{v}{n+1}\right|^3\right)\right\}, & \text{for} \quad 1 \leqq \left|\frac{v}{n+1}\right| \leqq 2 \\ 0 & \text{for} \quad \left|\frac{v}{n+1}\right| \geqq 2, \end{cases}$$

may be expressed as a Fourier-Stieltjes integral if one applies the result of Beurling [3, p. 349] below.

Lemma 8.1. *If $h(v)$ is absolutely continuous, and if both $h(v)$ and $h'(v)$ belong to $L_2(-\infty, \infty)$, then*

$$h(v) = \left(1/\sqrt{2\pi}\right) \int_{-\infty}^{\infty} e^{-ivx}\, dg(x),$$

where $g(x)$ is absolutely continuous.

It is readily seen that the conditions of this lemma are satisfied. By Theorem 5.2 the result follows.

According to Conjecture B, the saturation class of Jackson's singular integral in the space $L_1(-\infty, \infty)$ would be identical to that of Picard's integral.

For L_p-functions we have

Theorem 8.4. *Let $f \in L_p(-\infty, \infty)$, $1 < p \leqq 2$. For* Jackson's *integral the following are equivalent:*

1. $\|J_{2n}(x) - f(x)\|_{L_p} = O(1/n^2) \qquad (n \uparrow \infty)$;
2. *f and f' are absolutely continuous and f'' is in $L_p(-\infty, \infty)$.*

Furthermore, if $\|J_{2n}(x) - f(x)\|_{L_p} = o(1/n^2)$, then f is zero a.e.

§ 9. A Comparison with the Semi-group Method

We shall point out very briefly the connection between the Fourier transform methods of this paper and the semi-group methods. For this purpose we at first need several definitions.

Let X be a Banach space having elements f with norm $\|f\|$. Let

$$\mathfrak{S} = \{T(\zeta);\ 0 \leqq \zeta < \infty\}$$

be a family of bounded linear operators on X having the properties:

a) $T(\zeta_1 + \zeta_2) = T(\zeta_1)\, T(\zeta_2), \qquad T(0) f = f, \qquad \text{all } f \in X;$

b) $\lim_{\zeta \downarrow 0} \|T(\zeta) f - f\| = 0, \qquad \text{all } f \in X.$

The infinitesimal generator of $\mathfrak{S}$ is the operator A defined by

$$A f = \lim_{\zeta \downarrow 0} \frac{T(\zeta) - I}{\zeta} f$$

for all f, for which this limit exists (in the strong topology). The domain of A is denoted by $D(A)$.

Lemma 9.1. *Let $f \in X$.*

a) *If $\|T(\zeta) f - f\| = o(\zeta)$, $\zeta \downarrow 0$, then $Af = 0$ and $T(\zeta) f = f$, all $\zeta \geqq 0$.*

b) *If in addition the Banach space X is reflexive, then the following statements are equivalent:*

i) $\|T(\zeta) f - f\| = O(\zeta) \qquad (\zeta \downarrow 0)$,

ii) *f is in $D(A)$.*

For a proof one may consult [*8*] or [*28*, p. 326] and for a further generalization [*15*]. Part a) is due to Hille [*27*, p. 321], and for b) see Butzer [*8*].

It is this lemma which may be compared with Theorem 6.3. Indeed, the singular integral

$$J_\varrho(x) = \frac{\varrho}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x+u)\, k(\varrho u)\, du \tag{9.1}$$

defines a linear operator $S(\varrho)$ depending on the parameter $\varrho > 0$, mapping the functions f in the Banach space $L_\varrho(-\infty, \infty)$, $p \geqq 1$ into itself: $S(\varrho) f = J_\varrho$.

The operator $S(\varrho)$ has the property

$$[S(\varrho_1)\, S(\varrho_2)] f = S\left(\frac{\varrho_1\,\varrho_2}{\sqrt[\gamma]{\varrho_1^\gamma + \varrho_2^\gamma}}\right) f, \quad \text{every } f \in L_p(-\infty, \infty), \tag{9.2}$$

provided the kernel k satisfies the functional equation of CHAPMAN-KOLMOGOROFF

$$\frac{\varrho_1\,\varrho_2}{2\pi} \int_{-\infty}^{\infty} k\big(\varrho_1 (x-u)\big)\, k(\varrho_2 u)\, du = \frac{1}{\sqrt{2\pi}}\, \frac{\varrho_1\,\varrho_2}{\sqrt[\gamma]{\varrho_1^\gamma + \varrho_2^\gamma}}\, k\left(\frac{\varrho_1\,\varrho_2\, x}{\sqrt[\gamma]{\varrho_1^\gamma + \varrho_2^\gamma}}\right). \tag{9.3}$$

If we put $\varrho = 1/\gamma(\zeta)^{1/\gamma}$, $0 < \gamma \leqq 2$, then (9.2) assumes the form

$$\left[S\left(\frac{1}{\gamma \sqrt[\gamma]{\zeta_1}}\right) S\left(\frac{1}{\gamma \sqrt[\gamma]{\zeta_2}}\right)\right] f = S\left(\frac{1}{\gamma \sqrt[\gamma]{\zeta_1 + \zeta_2}}\right) f,$$

and this equation reveals the semi-group property of the operator $T(\zeta)$: $T(\zeta) f = S(\varrho) f$.

For example, if (9.1) is the Gauss-Weierstrass singular integral of § 7, then $\gamma = 2$, $\varrho = 1/2\sqrt{t}$, $\zeta = t$ and the equation (9.3) is of the form (compare [*27*, p. 403])

$$\frac{1}{4\sqrt{\pi^2 \zeta_1 \zeta_2}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-u)^2}{4\zeta_1}\right) \exp\left(\frac{-u^2}{4\zeta_2}\right) du = \frac{1}{2\sqrt{\pi(\zeta_1+\zeta_2)}} \exp\left(\frac{-x^2}{4(\zeta_1+\zeta_2)}\right).$$

For an example of a semi-group operator with $0 < \gamma < 2$ but $\gamma \neq 1$, see [*25*, p. *77*]. The kernel of this operator is the so-called *stable density of* PAUL LÉVY.

For further details in connection with the above we refer to [*6*, p. 69 *etc.*], [*28*, p. 566—571] and [*25*].

If $X = L_p(-\infty, \infty)$, $1 < p \leqq 2$, a comparison between the two methods is perhaps best revealed in the table of formal correspondences below.

Table

Semi-group Method	Fourier-transform Method
$\zeta > 0, \quad \zeta \downarrow 0$	$\varrho > 0, \quad \varrho \uparrow \infty\ (\zeta^{-1} = (\gamma\varrho)^\gamma)$
$[T(\zeta) f](x)$	$J_\varrho(x)$
$T(\zeta_1 + \zeta_2) = T(\zeta_1)\, T(\zeta_2)$	Formula (9.2)
$\Vert T(\zeta) f - f \Vert \to 0, \quad (\zeta \downarrow 0)$	$\Vert J_\varrho(x) - f(x) \Vert_{L_p} \to 0, \quad (\varrho \uparrow \infty)$
$\mathfrak{F}\left[\frac{T(\zeta) f - f}{\zeta}\right] \to \mathfrak{F}[A f]$ (compare [*28*], p. 370; p. 574—580)	$\varrho^\gamma [\hat{k}(v/\varrho) - 1]\, \hat{f}(v) \to -c\vert v\vert^\gamma \hat{f}(v)$
$D(A) = \{f \in X;\ A f$ exists and $\in X\}$	Saturation class of $J_\varrho(x) = \{f \in X\ c\vert v\vert^\gamma \hat{f}(v)$ is the Fourier transform of a function $g \in X\}$.

We have restricted the above discussion to the reflexive Banach space L_p, $1<p\leqq 2$, since Lemma 9.1 b) is thus restricted. For a version of Lemma 9.1 in the case $p=1$ we refer again to [*30*].

Of great importance in the Fourier-transform method is the condition (4.1). In a forthcoming paper [*13*], we intend to generalize this condition by expanding $\hat{k}(v/\varrho)$ in a Taylor polynomical about $v=0$. This would give results corresponding the to those in [*15*] for L_1-space.

References

[*1*] Alexits, G.: Sur l'ordre de grandeur de l'approximation d'une fonction par les moyennes de sa série de Fourier. Mat. és Fizikai Lapok **48**, 410—433 (1941).

[*2*] Aljančić, S.: Meilleure approximation et classes de saturation du procédé de Hölder dans les espaces C et L. Pub. Insti. Acad. Belgrade **12**, 109—124 (1958).

[*3*] Beurling, A.: Sur les intégrales de Fourier absolument convergentes et leur application à une transformation fonctionnelle. Neuvième Congrès Math. Scand., Helsingfors, 1938, 345—366.

[*4*] Bochner, S.: Vorlesungen über Fouriersche Integrale. Leipzig 1932. viii + 227 pp.

[*5*] Bochner, S., & K. Chandrasekharan: Fourier Transforms. Princeton 1949. 219 pp.

[*6*] Bochner, S.: Harmonic analysis and the Theory of Probability. Berkeley and Los Angeles 1955. viii + 176 pp.

[*7*] Buchwalter, H.: Saturation de certains procédés de summation. C. R. Acad. Sci. Paris **248**, 909—912 (1959).

[*8*] Butzer, P. L.: Über den Approximationsgrad des Identitätsoperators durch Halbgruppen von linearen Operatoren und Anwendungen auf die Theorie der singulären Integrale. Math. Annalen **133**, 410—425 (1957).

[*9*] Butzer, P. L.: Halbgruppen von linearen Operatoren und das Darstellungs- und Umkehrproblem für Laplace-Transformationen. Math. Annalen **134**, 154—166 (1957).

[*10*] Butzer, P. L.: Zur Frage der Saturationsklassen singulärer Integraloperatoren. Math. Z. **70**, 93—112 (1958).

[*11*] Butzer, P. L.: Representation and approximation of functions by general singular integrals. Nederl. Akad. Wetensch. Proc. Ser. A. 63 (= Indag. Math. **22**, 1—24 (1960).

[*12*] Butzer, P. L.: Sur le rôle de la transformation de Fourier dans quelques problèmes d'approximation. C. R. Acad. Sci. Paris **249**, 2467—2469 (1959).

[*13*] Butzer, P. L.: Asymptotic expansions of the solutions of the heat equation (in preparation).

[*14*] Butzer, P. L., & H. König: An application of Fourier-Stieltjes transforms in approximation theory. Arch. rational Mech. Anal. **5**, 416—419 (1960).

[*15*] Butzer, P. L., & H. G. Tillmann: Approximation theorems for semi-groups of bounded linear transformations. Math. Annalen **140**, 256—262 (1960).

[*16*] Churchill, R. V.: The operational calculus of Legendre transforms. J. Math. and Phys. **33**, 165—177 (1954).

[*17*] Cooper, J. L. B.: Fourier-Stieltjes integrals. Proc. London Math. Soc., Ser. II **51**, 265—284 (1950).

[*18*] Cooper, J. L. B.: The application of multiple Fourier transforms to the solution of partial differential equations. Quart. J. Math., Oxford Ser. (2) **1**, 122—135 (1950).

[*19*] Cramér, H.: On the representation of functions by certain Fourier integrals. Trans. Amer. Math. Soc. **46**, 190—201 (1939).

[*20*] Doetsch, G.: Integration von Differentialgleichungen vermittels der endlichen Fourier Transformationen. Math. Annalen **112**, 52—68 (1935).

[21] Doetsch, G.: Handbuch der Laplace-Transformation, Bd. I, II u. III. Basel 1950—1956.
[22] Erdélyi, A. et al.: Tables of Integral Transforms, Vol. I. New York 1954. xx + 391 pp.
[23] Favard, J.: Sur l'approximation des fonctions d'une variable réelle, Colloque d'Anal. Harmon. Publ. C. N. R. S. Paris **15**, 97—110 (1949).
[24] Favard, J.: Sur la saturation des procédés de sommation. J. de Math. **36**, 359—372 (1957).
[25] Feller, W.: On a generalization of Marcel Riesz' potentials and the semi-groups generated by them. Proc. R. Physiogr. Soc. Lund **21**, 73—81 (1952).
[26] Gonzáles Dominguez, A.: The representation of functions by Fourier integrals. Duke Math. J. **6**, 246—255 (1940).
[27] Hille, E.: Functional Analysis and Semi-groups. Amer. Math. Soc. Colloquium Publications **31**, New York 1948. xi+528 pp.
[28] Hille, E., & R. S. Phillips: Functional Analysis and Semi-groups. New York 1957 (revised edition). xii+808 pp.
[29] König, H.: Einige Eigenschaften der Fourier-Stieltjes Transformation. Arch. d. Math. **11** (1960) (in press).
[30] de Leeuw, K.: On the adjoint semi-group and some problems in the theory of approximation. Math. Z. **73** (1960) (in press).
[31] Sneddon, I. N.: Functional Analysis, in: Encyclopedia of Physics, edit. by S. Flügge, Vol. II, Mathematical Methods II, pp. 198—348. 1955.
[32] Sunouchi, G., & C. Watari: On determination of the class of saturation in the theory of approximation of functions. Proc. Japan Acad. **34**, No. 8, 477—481 (1958).
[33] Titchmarsh, E. C.: Introduction to the Theory of Fourier Integrals. Oxford 1937. viii+395 pp.
[34] Tureckiĭ, A. H.: On the saturation class in Hölder's method of summing Fourier series. Dokl. Akad. Nauk SSSR **121**, 980—983 (1958).
[35] Widder, D. V.: The Laplace Transform. Princeton 1941. x+406 pp.
[36] Zamansky, M.: Classes de saturation de certains procédés d'approximation. Ann. sci. École norm. sup. fasc. **1**, 19—93 (1949).
[37] Zygmund, A.: Trigonometrical Series, 2 Vols. Cambridge 1959 (revised edition).

Department of Mathematics
The Technical University
Aachen, Germany

(Received April 4, 1960)

An Application of Fourier-Stieltjes Transforms in Approximation Theory

PAUL L. BUTZER & HEINZ KÖNIG

Communicated by C. MÜLLER

1. Introduction

It is the aim of this paper to continue the investigations on the degree of approximation of functions by general singular integrals started by one of the authors [*3, 4*]. The basic idea is to treat these problems by the so-called Fourier-transform method. In the present paper we give a very short proof of [*4*] Theorem 4.2 (Theorem 1 below) using a representation theorem of BOCHNER [*2*] which also leads to another proof of [*4*] Theorem 5.1 (Theorem 2 below). A generalization of BOCHNER's theorem due to R. S. PHILLIPS [*6*] is then applied to obtain a certain converse to the second theorem (Theorem 4 below). The conditions to be imposed upon the kernel will be investigated in [*5*].

A function P defined on $-\infty < v < \infty$ is said to be of class $\mathfrak{F}$ if it is the Fourier transform

$$P(v) = \hat{f}(v) = (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} e^{-ixv} f(x)\, dx$$

of a function $f \in L_1(-\infty, \infty)$. Also P is of class $\mathfrak{F}\mathfrak{S}$ if it is the Fourier-Stieltjes transform

$$P(v) = \check{g}(v) = (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} e^{-ixv}\, dg(x)$$

of a function g of bounded variation on $-\infty < x < \infty$. For $P \in \mathfrak{F}\mathfrak{S}$ we introduce the norm

$$\|P\| = \|\check{g}\| = (1/\sqrt{2\pi})\, [\operatorname{Var} g]_{-\infty}^{\infty}. \tag{1.1}$$

If $P \in \mathfrak{F}$, we have

$$\|P\| = \|\hat{f}\| = (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} |f(x)|\, dx = \|f\|_1. \tag{1.2}$$

If $P \in \mathfrak{F}\mathfrak{S}$, it follows immediately that

$$\left| \sum_{j=1}^{r} c_j P(v_j) \right| \le \|P\| \left\| \sum_{j=1}^{r} c_j e^{-iv_j x} \right\|_{\infty} \tag{1.3}$$

for every set of complex $c_1, \ldots, c_r$ and real $v_1, \ldots, v_r$. The converse to the latter fact is given by the following theorems.

Bochner Theorem. *If P is continuous on $-\infty < v < \infty$ and if there exists an $M > 0$ such that*

$$\left| \sum_{j=1}^{r} c_j P(v_j) \right| \le M \left\| \sum_{j=1}^{r} c_j e^{-iv_j x} \right\|_{\infty} \tag{1.4}$$

for all complex $c_1, \ldots, c_r$ and real $v_1, \ldots, v_r$, $r = 1, 2, \ldots$, then $P \in \mathfrak{F}\mathfrak{S}$ and $\|P\| \le M$.

Phillips Theorem. *If P is measurable on $-\infty<v<\infty$ and if there exists an $M>0$ such that (1.4) is satisfied, then P is almost everywhere equal to a function $Q\in\mathfrak{F}\mathfrak{S}$ and $\|Q\|\leqq M$.*

2. The Singular Integral

We consider the singular integral

(2.1)
$$\begin{aligned} J_\varrho(x) &= (\varrho/\sqrt{2\pi}) \int_{-\infty}^{\infty} f(x-u)\, k(\varrho u)\, du \\ &= (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} f\left(x-\frac{u}{\varrho}\right) k(u)\, du \qquad (\varrho>0) \end{aligned}$$

for $f\in L_1(-\infty,\infty)$. The kernel k is an arbitrary function in $L_1(-\infty,\infty)$ with

$$(1/\sqrt{2\pi}) \int_{-\infty}^{\infty} k(u)\, du = 1\,.$$

Clearly $J_\varrho(x)$ exists for almost all x, and $J_\varrho\in L_1(-\infty,\infty)$ with $\|J_\varrho\|_1\leqq\|f\|_1\,\|k\|_1$. Also $\|J_\varrho-f\|_1\to 0$ for $\varrho\to\infty$.

Assume that $\gamma>0$. We denote by $\mathfrak{A}(\gamma)$ the class of functions $f\in L_1(-\infty,\infty)$ for which

$$\|J_\varrho-f\|_1 = O(\varrho^{-\gamma}) \qquad (\varrho\to\infty)\,.$$

For $\gamma>0$ we introduce the function

$$H^0_\gamma(v) = \left\{\begin{array}{ll} \dfrac{1-\hat{k}(v)}{|v|^\gamma} & (v\neq 0) \\ 0 & (v=0) \end{array}\right\};$$

note that $\hat{k}(0)=1$. If $H^0_\gamma(v)$ tends to a limit c for $v\to 0$, then we put

$$H_\gamma(v) = \left\{\begin{array}{ll} \dfrac{1-\hat{k}(v)}{|v|^\gamma} & (v\neq 0) \\ c & (v=0) \end{array}\right\}.$$

The convolution theorem for Fourier transforms applied to (2.1) gives

$$\hat{J}_\varrho(v) = \hat{f}(v)\,\hat{k}(v/\varrho) \qquad (\varrho>0)\,.$$

Thus the function

$$\varrho^\gamma \hat{f}(v)\,\bigl(1-\hat{k}(v/\varrho)\bigr) = |v|^\gamma \hat{f}(v)\, H^0_\gamma(v/\varrho)$$

is the Fourier transform of $\varrho^\gamma(f-J_\varrho)$ so that by (1.2)

$$\|\,|v|^\gamma \hat{f}(v)\, H^0_\gamma(v/\varrho)\| = \varrho^\gamma\|J_\varrho-f\|_1\,.$$

According to (1.3) we have from the Bochner theorem: The function $f\in L_1(-\infty,\infty)$ belongs to $\mathfrak{A}(\gamma)$ if and only if there is an $M>0$ such that

(2.2)
$$\left|\sum_{j=1}^{r} c_j\,|v_j|^\gamma \hat{f}(v_j)\, H^0_\gamma(v_j/\varrho)\right| \leqq M\left\|\sum_{j=1}^{r} c_j\, e^{-i v_j x}\right\|_\infty$$

for all $c_1,\ldots,c_r$ and $v_1,\ldots,v_r$ and every $\varrho>0$.

3. The Direct and Inverse Theorems

The Bochner-Phillips theorem provides us with a unified method of proof of the subsequent results. However this method seems to be restricted to L_1-space.

Theorem 1. *Assume that* $H_\gamma^0(v) \to c \neq 0$ *for* $v \to 0$. *Then* $f \in \mathfrak{A}(\gamma)$ *implies that* $|v|^\gamma \hat{f}(v) \in \mathfrak{F}\mathfrak{S}$.

Proof. We pass to the limit $\varrho \to \infty$ in (2.2) and apply the Bochner theorem to the function $P(v) = |v|^\gamma \hat{f}(v)$.

Theorem 2. *Assume that* H_γ *exists and is in* $\mathfrak{F}\mathfrak{S}$. *Then every* $f \in L_1(-\infty, \infty)$ *such that* $|v|^\gamma \hat{f}(v) \in \mathfrak{F}\mathfrak{S}$ *belongs to* $\mathfrak{A}(\gamma)$.

Proof. The application of (1.3) to the function $|v|^\gamma \hat{f}(v)$ gives

$$\begin{aligned}\left|\sum_{j=1}^{r} c_j |v_j|^\gamma \hat{f}(v_j)\, H_\gamma^0(v_j/\varrho)\right| &= \left|\sum_{j=1}^{r} c_j |v_j|^\gamma \hat{f}(v_j)\, H_\gamma(v_j/\varrho)\right|, \\ &\leqq M_1 \left\|\sum_{j=1}^{r} c_j e^{-i v_j x} H_\gamma(v_j/\varrho)\right\|_\infty,\end{aligned} \tag{3.1}$$

where $M_1 = \| |v|^\gamma \hat{f}(v) \|$. A second application of (1.3) to the function $H_\gamma(v/\varrho)$ yields

$$\begin{aligned}\left|\sum_{j=1}^{r} c_j e^{-i v_j t} H_\gamma(v_j/\varrho)\right| &\leqq M_2 \left\|\sum_{j=1}^{r} c_j e^{-i v_j t} e^{-i v_j x/\varrho}\right\|_\infty, \\ &= M_2 \left\|\sum_{j=1}^{r} c_j e^{-i v_j x}\right\|_\infty\end{aligned} \tag{3.2}$$

for every t. Here $M_2 = \|H_\gamma(v)\|$ since for $H_\gamma(v) = \check{g}(v)$ we have

$$\begin{aligned}\|\check{g}(v/\varrho)\| &= (1/\sqrt{2\pi})\,[\operatorname{Var} g(\varrho x)]_{-\infty}^{\infty} \\ &= (1/\sqrt{2\pi})\,[\operatorname{Var} g(x)]_{-\infty}^{\infty} = \|\check{g}(v)\|.\end{aligned}$$

Combining the inequalities (3.1) and (3.2), we obtain (2.2) with $M = M_1 M_2$, establishing the theorem.

The purpose of the following considerations is to obtain a converse to Theorem 2.

Theorem 3. *Let* $\mathfrak{A}(\gamma)$ *contain a function* $f \in L_1(-\infty, \infty)$ *such that* $|v|^\gamma \hat{f}(v) \to R \neq 0$ *for* $|v| \to \infty$. *Then* H_γ *exists and is in* $\mathfrak{F}\mathfrak{S}$.

Proof. Substituting v_j by $v_j \varrho$ in (2.2), we have

$$\left|\sum_{j=1}^{r} c_j |v_j \varrho|^\gamma \hat{f}(v_j \varrho)\, H_\gamma^0(v_j)\right| \leqq M \left\|\sum_{j=1}^{r} c_j e^{-i v_j x}\right\|_\infty.$$

Now we let $\varrho \to \infty$. Since $H_\gamma^0(0) = 0$,

$$\left|\sum_{j=1}^{r} c_j H_\gamma^0(v_j)\right| \leqq \frac{M}{|R|} \left\|\sum_{j=1}^{r} c_j e^{-i v_j x}\right\|_\infty,$$

and the desired result follows from the theorem of PHILLIPS.

Lemma. *We have* $(1+|v|^{\gamma})^{-1}\in\mathfrak{F}$ *for every* $\gamma>0$.

Proof. In the case when $0<\gamma\leqq 1$, the function under consideration is monotone decreasing and convex in $v\geqq 0$ and thus, according to TITCHMARSH [*7*, Theorem 124, p. 170], belongs to $\mathfrak{F}$. For $\gamma>\frac{1}{2}$, the function is in $L_2(-\infty,\infty)$ and is furthermore absolutely continuous with derivative in $L_2(-\infty,\infty)$. A theorem of BEURLING [*1*, p. 349] gives the result.

Theorem 3 combined with the lemma gives the desired converse.

Theorem 4. *Suppose that every* $f\in L_1(-\infty,\infty)$ *with* $|v|^{\gamma}\hat{f}(v)\in\mathfrak{F}\mathfrak{S}$ *belongs to* $\mathfrak{A}(\gamma)$. *Then* H_γ *exists and is in* $\mathfrak{F}\mathfrak{S}$.

Proof. The function $f\in L_1(-\infty,\infty)$ with Fourier transform $\hat{f}(v)=(1+|v|^{\gamma})^{-1}$ belongs to $\mathfrak{A}(\gamma)$ since $|v|^{\gamma}\hat{f}(v)=1-\hat{f}(v)$ is in $\mathfrak{F}\mathfrak{S}$. But $|v|^{\gamma}\hat{f}(v)\to 1$ for $|v|\to\infty$, so that the result follows from Theorem 3.

We conclude by remarking that for many kernels only the case $0<\gamma\leqq 2$ is of interest.

Theorem 5. *Let* $k(u)$ *be even and non-negative. Then* $\mathfrak{A}(\gamma)$ *contains only the null function in case* $\gamma>2$.

Proof. From the equation

$$H_\gamma^0(x)=\frac{1-\hat{k}(x)}{|x|^{\gamma}}=\frac{1}{|x|^{\gamma-2}}\,\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{1-\cos xu}{x^2}\,k(u)\,du$$

it follows that $H_\gamma^0(x)\to\infty$ for $x\to 0$, since $\gamma>2$. Now for any function $f\in\mathfrak{A}(\gamma)$, applying (2.2) for $r=1$, we have

$$|v|^{\gamma}\,|\hat{f}(v)|\,|H_\gamma^0(v/\varrho)|\leqq M$$

for $\varrho>0$ and all v. Letting $\varrho\to\infty$, it follows that $\hat{f}(v)=0$ for $v\neq 0$. The uniqueness theorem shows that f is the null function.

References

[*1*] BEURLING, A.: Sur les Intégrales de Fourier absolument convergentes et leur Application à une Transformation fonctionnelle. Neuvième Congrès des Mathématiciens Scandinaves, Helsingfors 1938, 345—366.

[*2*] BOCHNER, S.: A Theorem on Fourier-Stieltjes Integrals. Bull. Amer. Math. Soc. **40**, 271—276 (1934).

[*3*] BUTZER, P. L.: Sur le rôle de la Transformation de Fourier dans quelques problèmes d'approximation. C. R. Acad. Sci. Paris **249**, 2467—2469 (1959).

[*4*] BUTZER, P. L.: Fourier-Transform Methods in the Theory of Approximation. Arch. rational Mech. Anal. **5**, 390—415 (1960).

[*5*] KÖNIG, H.: Einige Eigenschaften der Fourier-Stieltjes Transformation. (To appear in Arch. Math.)

[*6*] PHILLIPS, R. S.: On Fourier-Stieltjes Integrals. Trans. Amer. Math. Soc. **69**, 312—323 (1950).

[*7*] TITCHMARSH, E. C.: Introduction to the Theory of Fourier Integrals, second ed. Oxford 1948.

Department of Mathematics
The Technical University
Aachen, Germany

(Received April 4, 1960)

Eine Anwendung des Nirenbergschen Maximumprinzips für parabolische Differentialgleichungen in der Grenzschichttheorie

W. VELTE

Vorgelegt von H. GÖRTLER

Von K. NICKEL [*6*] wurde gezeigt, daß man auch ohne explizite Integration der Prandtlschen Grenzschichtgleichungen bereits eine ganze Anzahl allgemeiner Aussagen über das mögliche Verhalten von ebenen Grenzschichtströmungen erhalten kann, z.B. über Entstehen und Anwachsen von Übergeschwindigkeiten, über die Maxima und Minima der Schubspannung, und darauf aufbauend auch über Extrema und Wendepunkt der Grenzschichtprofile.

Diese Ergebnisse hat NICKEL mit Hilfe eines Abschätzungssatzes von NAGUMO-WESTPHAL [*5*], [*9*] gewonnen, der auch schon von H. GÖRTLER [*1*] in der Grenzschichttheorie verwendet wurde. Um diesen Satz anwenden zu können, mußte NICKEL zuvor die Prandtlschen Grenzschichtgleichungen für die zwei Komponenten der Geschwindigkeit auf eine einzige partielle Differentialgleichung von nur einer gesuchten Funktion transformieren, nämlich auf die v. Misessche bzw. die Croccosche Differentialgleichung. Eine solche Transformation ist aber nur für ebene Strömungen möglich. Daran liegt es, daß NICKEL seine Ergebnisse nur im 2-dimensionalen Fall erhalten konnte.

Im folgenden soll gezeigt werden, daß es noch einen anderen und zudem sehr einfachen Zugang zu den grundlegenden Nickelschen Sätzen gibt, bei dem eine Transformation der Prandtlschen Differentialgleichungen nicht erforderlich ist. Die eigentliche Bedeutung der neuen Methode besteht aber darin, daß sich in genau der gleichen Weise auch viel allgemeinere Fälle behandeln lassen, nämlich die 3-dimensionalen Grenzschichtgleichungen, und zwar stationär wie instationär. Außerdem können an der Wand weitgehend willkürliche Werte für die Komponenten der Geschwindigkeit vorgegeben werden, d.h. es darf in sehr allgemeiner Weise abgesaugt bzw. ausgeblasen werden. Dagegen wird die einschränkende Voraussetzung einer rotationsfreien Außenströmung beibehalten. Es lassen sich dann für die 3-dimensionalen Strömungen Aussagen gewinnen über das Anwachsen der beiden Geschwindigkeitskomponenten in Hauptströmungsrichtung, in denen die entsprechenden Nickelschen Sätze als Sonderfälle enthalten sind. Dabei ergeben sich sogar noch einige Verschärfungen.

Aus dem Institut für Angewandte Mathematik und Mechanik der DVL an der Universität Freiburg i. Br. Diese Untersuchung wurde gefördert vom Wirtschaftsministerium des Landes Baden-Württemberg.

Als mathematisches Hilfsmittel tritt jetzt an die Stelle des Nagumo-Westphalschen Satzes für nichtlineare parabolische Differentialgleichungen das Maximumprinzip von NIRENBERG [7] für lineare parabolische Differentialoperatoren 2. Ordnung, d.h. das Analogon des bekannten Maximumprinzips von E. HOPF [2] für elliptische Differentialoperatoren.

Wir beschränken uns in der vorliegenden Arbeit darauf, die inkompressiblen Grenzschichtgleichungen zu diskutieren. Es sei jedoch vermerkt, daß die Methode weiter trägt und sogar die kompressiblen wie auch die Grenzschichtgleichungen in krummlinigen Koordinaten in ähnlicher Weise zu behandeln gestattet, worüber in Kürze berichtet wird.

1. Maximumprinzip bei linearen parabolischen Operatoren

Im folgenden bezeichnet G stets eine offene, nicht notwendig beschränkte, zusammenhängende Punktmenge im Raum der Variablen $x_1, x_2, \ldots, x_n$ und $t_1, t_2, \ldots, t_m$. Wir schreiben auch kurz $x = (x_1, x_2, \ldots, x_n)$ und $|x|^2 = \sum x_i^2$ $(i = 1, 2, \ldots, n)$. Analog sind t bzw. $|t|^2$ zu lesen.

Es werden lineare Differentialoperatoren der Form

(1.1)
$$L \equiv \sum a_{ij}(x,t) \frac{\partial^2}{\partial x_i \partial x_j} + \sum b_i(x,t) \frac{\partial}{\partial x_i} + \sum c_k(x,t) \frac{\partial}{\partial t_k}$$
$$(i,j = 1, 2, \ldots, n;\;\; k = 1, 2, \ldots, m)$$

betrachtet. Die Koeffizienten seien in G als stetige Funktionen erklärt, und die quadratische Form

(1.2)
$$Q \equiv \sum a_{ij}(x,t)\, \xi_i \xi_j \qquad (i,j = 1, 2, \ldots, n)$$

sei in jedem Punkt (x, t) aus G positiv definit.

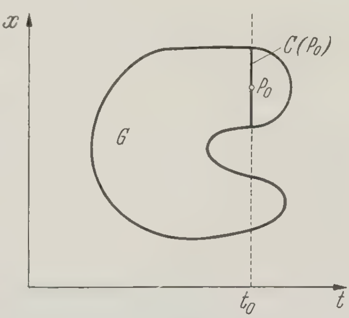

Abb. 1. Zur Definition von $C(P_0)$

Alle Funktionen, auf die im folgenden der Operator (1.1) angewandt wird, werden in G stetig vorausgesetzt, desgleichen ihre ersten Ableitungen nach x und t sowie die zweiten Ableitungen nach x.

Zur Formulierung des Maximumprinzips übernehmen wir aus [7] folgende Definition: Es sei P_0 ein Punkt aus G. Man betrachte die n-dimensionale Hyperebene $t_k = \text{const}$ $(k = 1, 2, \ldots, m)$ durch den Punkt P_0 und bilde ihren Durchschnitt mit G. Dann bezeichnet $C(P_0)$ die zusammenhängende Komponente des Durchschnittes, die P_0 enthält. (Vgl. Abb. 1.)

Satz 1.1 (NIRENBERG). *Die Funktion $f(x,t)$ genüge in ganz G der Ungleichung $L(f) \geqq 0$ $(\leqq 0)$. Wenn dann $f(x,t)$ im Punkte $P_0 \in G$ ihr Maximum (Minimum) annimmt, dann ist $f(P) = f(P_0)$ für alle $P \in C(P_0)$.*

Dieser Satz läßt sich auf gewisse Randpunkte von G ausdehnen. Bei NIRENBERG wird es in der Form nicht ausgesprochen, ergibt sich aber aus seinen Beweisen mit einer geringen Modifikation. Diese Ergänzung für Randpunkte ist

in Satz 1.2 formuliert. Im Anhang wird noch angegeben, wie man den Nirenbergschen Beweis abzuändern hat [1].

Es werde ein Gebiet G betrachtet, das ein zusammenhängendes Randstück R' besitzt mit folgenden Eigenschaften (vgl. Abb. 2):

(a) Alle Punkte von R' liegen auf mindestens einer der m Hyperebenen $t_k = \text{const}$ $(k=1, 2, \ldots, m)$.

(b) Zu jedem Punkt $P_0 \in R'$ mit den Koordinaten (x_0, t_0) gibt es ein $r > 0$, so daß die Menge H

$$H(P_0): \quad |x-x_0|^2 + |t-t_0|^2 < r^2, \qquad t_k \leqq t_{k0}$$

ganz zu $G+R'$ gehört.

(Im Falle $m=1$ ist H eine Halbkugel, im allgemeinen Fall Segment einer Hyperkugel.) Die Stetigkeit und Differenzierbarkeit der Funktionen $f(x, t)$ bzw. der Koeffizienten von (1.1) wird für Satz 1.2 auf der Punktmenge $G+R'$ vorausgesetzt. Die quadratische Form (1.2) sei auch auf R' positiv definit. Dehnt man die Definition von $C(P_0)$ auf die Menge $G+R'$ aus, dann lautet die Ergänzung zu Satz 1.1 bezüglich des Randes R':

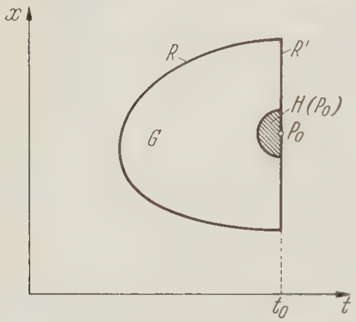

Abb. 2. Zur Definition des Randes R'

Satz 1.2. *Die Funktion $f(x, t)$ genüge in $G+R'$ der Ungleichung $L(f) \geqq 0$ $(\leqq 0)$. Ferner sei auf R' stets $c_k(x, t) \leqq 0$ $(k=1, 2, \ldots, m)$. Wenn dann $f(x,t)$ im Punkt $P_0 \in R'$ ihr Maximum (Minimum) annimmt, dann ist $f(P) = f(P_0)$ für alle Punkte $P \in C(P_0)$.*

2. Stationäre Grenzschichtströmungen

Im folgenden werden Lösungen der 3-dimensionalen stationären Grenzschichtgleichungen

$$\begin{aligned} u\,u_x + v\,u_y + w\,u_z &= -\frac{1}{\varrho}\,p_x + \nu\,u_{yy} \\ u\,w_x + v\,w_y + w\,w_z &= -\frac{1}{\varrho}\,p_z + \nu\,w_{yy} \end{aligned} \tag{2.1}$$

$$u_x + v_y + w_z = 0 \tag{2.2}$$

betrachtet. Hierin bezeichnen x, y, z kartesische Koordinaten. Wie üblich in der Grenzschichttheorie ist y die Koordinate senkrecht zur Wand. Ferner ist ν die kinematische Zähigkeit $(\nu > 0)$ und ϱ die als konstant betrachtete Dichte der Strömung. Die Druckverteilung p ergibt sich aus der rotationsfrei vorausgesetzten Außenströmung $U = U(x, z) > 0$, $W = W(x, z) > 0$ mittels $\frac{1}{\varrho}\,p + \frac{1}{2}(U^2 + W^2) = \text{const}$. Die Funktionen U, W und die Komponenten u, v, w der Geschwindigkeit sollen so oft stetig differenzierbar sein, wie es in den Differentialgleichungen (2.1)

[1] Für elliptische Operatoren hat schon E. HOPF [3] Randmaxima untersucht (vergl. auch MIRANDA [4]). Für elliptisch-parabolische Operatoren findet man entsprechende Sätze z.B. bei C. PUCCI [8]. Die in Satz 1.2 formulierte Ergänzung im Anschluß an NIRENBERG [7] ist speziell auf unsere Anwendungen zugeschnitten.

und (2.2) vorkommt. Außerdem soll die äußere Randbedingung

(2.3) $$\lim_{y\to\infty}(u - U) = \lim_{y\to\infty}(w - W) = 0$$

erfüllt sein.

Neben den Komponenten u, w selbst werden wir noch die skalare Größe $u^2+w^2-U^2-W^2$ betrachten. Sie stellt (bis auf einen konstanten Faktor) die Differenz der kinetischen Energie eines Volumenelementes gegenüber der kinetischen Energie der Außenströmung dar. Zu jeder Stelle (x, z) der Wand $(y=0)$ kann man sich dann diese Größe in Abhängigkeit von y auftragen. An die Betrachtung dieser ,,Energieprofile" knüpft sich eine wichtige Maximumeigenschaft, die eine a priori Abschätzung von u^2+w^2 gestattet.

Es werde nun eine bestimmte Lösung u, v, w der Gleichungen (2.1) und (2.2) ins Auge gefaßt. Mit diesen Funktionen erklären wir einen Operator L durch

(2.4) $$L \equiv \nu\frac{\partial^2}{\partial y^2} - v\frac{\partial}{\partial y} - u\frac{\partial}{\partial x} - w\frac{\partial}{\partial z}.$$

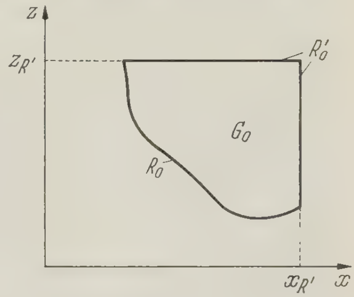

Abb. 3. Zur Definition von G_0, R_0, R_0'

Hierin ist y die ,,elliptische" Variable $(n=1)$, und x, z sind die ,,parabolischen" Variablen $(m=2)$. Die quadratische Form (1.2) lautet hier einfach $Q=\nu\xi^2$, ist also positiv definit. Wir wenden diesen Operator der Reihe nach an auf die Funktionen $u, w, u^2+w^2-U^2-W^2$. Unter Benutzung von (2.1) und (2.2) erhält man

(2.5) $$L(u) = -(U U_x + W W_x) \quad \text{bzw.} \quad L(w) = -(U U_z + W W_z),$$

(2.6) $$L(u^2+w^2-U^2-W^2) = 2\nu(u_y)^2 + 2\nu(w_y)^2 \geqq 0.$$

Hinzu kommt noch im ebenen Fall $\big(w=W=0,\ U=U(x)$, alle Funktionen von z unabhängig$\big)$, wenn zusätzlich Existenz und Stetigkeit von u_{yyy}, $u_{xy}=u_{yx}$ vorausgesetzt werden,

(2.7) $$L(u_y) = 0.$$

Die Anwendung des Maximumprinzips liegt nun auf der Hand. An sich lassen sich Sätze für sehr allgemeine Gebiete G formulieren. Für die Diskussion von Grenzschichten sind aber vor allem zylindrische Gebiete von Interesse, auf die wir uns hier beschränken wollen.

Es sei zunächst G_0 ein offenes, zusammenhängendes, beschränktes Gebiet in der x, z-Ebene, das gemäß Abb. 3 von zwei Geradenstücken mit $x=x_{R'}$ bzw. $z=z_{R'}$ begrenzt wird, so daß für alle Punkte (x, z) aus G_0 stets $x<x_{R'}$, $z<z_{R'}$ gilt. Die Bedeutung der Ränder R_0 (abgeschlossen) und R_0' (offen) geht aus Abb. 3 hervor.

Mit G werde dann der Zylinder $0<y<\infty$ bezeichnet, dessen Punkte sich längs der y-Achse auf G_0 projizieren. R' bezeichnet die Randpunkte von G mit $0<y<\infty$, die sich auf R_0' projizieren. R ist die Gesamtheit aller übrigen Randpunkte von G. Es ist unmittelbar klar, daß der Rand R' die Eigenschaften (a) und (b) aus Abschnitt 1 besitzt. Es wird nun eine in $G+R+R'$ stetige Lösung u, v, w der Gleichungen (2.1) und (2.2) betrachtet, deren Ableitungen, soweit sie in den Gleichungen auftreten, in $G+R'$ stetig seien.

In den folgenden Sätzen wird an der Wand ($y=0$) über die Komponente v in y-Richtung keine einschränkende Voraussetzung gemacht. Dagegen sind die Komponenten u, w in Hauptströmungsrichtung dort nicht ganz willkürlich.

Satz 2.1. *Es sei $u \geqq 0$, $w \geqq 0$ in ganz $G+R+R'$. Wenn außerdem $UU_x+WW_x \geqq 0$ ist, dann gilt sogar $u>0$ in ganz $G+R'$.* (*Das gleiche gilt für w, wenn $UU_z+WW_z \geqq 0$.*)

Beweis. Nach (2.5) ist $L(u) \leqq 0$, so daß man das Minimumprinzip in $G+R'$ anwenden kann. Würde nun im Punkte $P_0 \in G+R'$ der Wert $u=0$ angenommen, dann hätte die Funktion u, von der durchweg $u \geqq 0$ gilt, dort ihr Minimum erreicht. Nach Satz 1.1 bzw. Satz 1.2 müßte dann auch $u=0$ sein auf ganz $C(P_0)$, d.h. auf der Halbgeraden $x=x_0$, $z=z_0$, $0<y<\infty$, im Widerspruch zur äußeren Randbedingung $u \to U>0$ für $y \to \infty$.

Im folgenden werden wir die äußere Randbedingung in etwas strengerer Form voraussetzen, nämlich mit

$$\lim_{y\to\infty}(u-U) = \lim_{y\to\infty}(w-W) = 0 \qquad \textit{gleichmäßig} \text{ in } x \text{ und } z. \tag{2.8}$$

Das heißt nach Vorgabe eines positiven ε ist es möglich das y_2 so zu wählen, daß für $y \geqq y_2$ und beliebiges (x, z) aus $G_0+R_0+R_0'$ immer gilt: $|u-U| \leqq \varepsilon$, $|w-W| \leqq \varepsilon$ [2]. Aus (2.8) ergibt sich dann auch

$$\lim_{y\to\infty}(u^2-U^2) = \lim_{y\to\infty}(w^2-W^2) = 0 \qquad \textit{gleichmäßig} \text{ in } x \text{ und } z. \tag{2.9}$$

Ferner führen wir noch ein:

$$\bar{u} = \sup_{P\in R} u, \qquad \bar{w} = \sup_{P\in R} w, \qquad M = \sup_{P\in R}(u^2+w^2-U^2-W^2). \tag{2.10}$$

Es ist immer $M \geqq 0$ wegen $u^2+w^2-U^2-W^2 \to 0$ für $y \to \infty$. Ist $M>0$, so hat man „Übergeschwindigkeiten" auf R, d.h. Stellen, in denen $u-U>0$ oder $w-W>0$ ist.

Satz 2.2. (a) *Es sei $u \geqq 0$, $w \geqq 0$ auf R'. Die äußere Randbedingung sei für u in der strengeren Form* (2.8) *erfüllt. Ferner sei $0<U \leqq \bar{u}$, $UU_x+WW_x \leqq 0$ in $G+R'$. Dann ist $u \leqq \bar{u}$ in ganz $G+R'$.*

(b) *Wenn an der Wand zusätzlich $u<\bar{u}$ erfüllt ist, dann gilt sogar $u<\bar{u}$ in ganz $G+R'$.*

(Das gleiche gilt für w, wenn die äußere Randbedingung für w in der strengeren Form erfüllt ist, und wenn $0<W \leqq \bar{w}$, $UU_z+WW_z \leqq 0$ ist.)

Kurz gesagt bedeutet der Satz folgendes: Wenn auf dem hinteren Rande R' des durchströmten Gebietes nirgends Rückströmung vorliegt, und wenn die kinetische Energie der Außenströmung nirgends zunimmt, dann kann eine auf R vorhandene Übergeschwindigkeit nicht stärker anwachsen als bis zu $\bar{u}$ bzw. $\bar{w}$.

[2] Man hätte diese Forderung auch anders fassen können, wie es NICKEL im ebenen Fall getan hat: Nach der Transformation

$$y = \frac{\eta}{1-\eta}$$

sollen die Funktionen $u(x, y(\eta), z)$, $v(x, y(\eta), z)$ in dem *endlichen* Gebiet $\bar{G}$ mit $(x, z) \in G_0+R_0+R_0'$ und $0 \leqq \eta \leqq 1$ stetig sein, und es soll gelten

$$\lim_{\substack{\eta\to 1\\ P\in\bar{G}}} u(x, y(\eta), z) = U(x, z) \quad \text{bzw.} \quad \lim_{\substack{\eta\to 1\\ P\in\bar{G}}} w(x, y(\eta), z) = W(x, z).$$

Liegt auf R nirgends eine Übergeschwindigkeit vor, dann kann auch keine mehr entstehen*.

Beweis. (a) Angenommen, es gäbe Punkte mit $u > \bar{u}$, z.B. sei $u = \bar{u} + \alpha$ $(\alpha > 0)$ in $P \in G + R'$. Dann kann man wegen (2.8) ein y_2 so wählen, daß für alle $y \geqq y_2$ immer $u - U \leqq \frac{\alpha}{2}$ gilt, und daher wegen $U \leqq \bar{u}$ auch immer $u \leqq \bar{u} + \frac{\alpha}{2}$.

Jetzt wird der bei $y = y_2$ abgeschnittene Zylinder betrachtet. Es bezeichne G_* (bzw. R'_*) die Gesamtheit der Punkte von G (bzw. von R') mit $0 < y < y_2$. R_* bezeichne die Gesamtheit der Randpunkte von G_*, die nicht zu R'_* gehören. $G_* + R_* + R'_*$ ist eine beschränkte und abgeschlossene Punktmenge, auf der die stetige Funktion u ihr Maximum annimmt. Und zwar ist $\operatorname{Max}(u) \geqq \bar{u} + \alpha$ und wird in einem Punkt P_0 aus $G + R'$ angenommen. Da nun nach (2.5) immer $L(u) \geqq 0$ ist, kann man das Maximumprinzip anwenden: Das Maximum wird auf ganz $C_*(P_0)$ angenommen, d.h. auf dem offenen Geradenstück $x = x_0$, $z = z_0$, $0 < y < y_2$ und wegen der Stetigkeit von u auch in den Endpunkten, was aber im Widerspruch steht zu $u \leqq \bar{u} + \frac{\alpha}{2}$ für $y \geqq y_2$. Es ist daher $u \leqq \bar{u}$ in ganz $G + R'$.

(b) Würde der Wert $u = \bar{u}$ in $P_0 \in G + R'$ angenommen, so hätte die Funktion u, von der ja $u \leqq \bar{u}$ gilt, dort ihr Maximum erreicht. Dann wäre wieder $u = \bar{u}$ auf ganz $C(P_0)$, d.h. auf $x = x_0$, $z = z_0$, $0 < y < \infty$ mit Einschluß des Endpunktes $y = 0$, im Widerspruch zur Voraussetzung $u < \bar{u}$ an der Wand.

Der nächste Satz bezieht sich auf die kinetische Energie und enthält eine a priori Abschätzung für $u^2 + w^2$. Mit der Konstanten M aus (2.10) lautet er

Satz 2.3. (a) *Es sei $u \geqq 0$, $w \geqq 0$ auf dem Rand R'. Die äußere Randbedingung sei für u und w in der strengeren Form* (2.8) *erfüllt. Dann ist*

$$u^2 + w^2 \leqq U^2 + W^2 + M \quad \text{in} \quad G + R'.$$

(b) *Wenn die Ungleichung zusätzlich an der Wand mit dem Zeichen $<$ erfüllt ist, dann gilt sogar*

$$u^2 + w^2 < U^2 + W^2 + M \quad \text{in} \quad G + R'.$$

Beweis. (a) Wir schreiben $f = u^2 + w^2 - U^2 - W^2$. Angenommen, es gäbe einen Punkt $P \in G + R'$, in dem $f > M$ ist. Es sei etwa $f(P) = M + \alpha\,(\alpha > 0)$. Nach (2.9) kann man dann ein y_2 so wählen, daß für alle $y \geqq y_2$ gleichmäßig in x, z stets $|f| \leqq M + \frac{\alpha}{2}$ gilt. Nun wird wieder das bei y_2 abgeschnittene Gebiet G_* betrachtet. f nimmt auf $G_* + R_* + R'_*$ das Maximum an: $\operatorname{Max}(f) \geqq M + \alpha$. Auf dem Rande R_* ist aber $f \leqq M + \frac{\alpha}{2}$. Also muß das Maximum in einem Punkt $P_0 \in G + R'$ angenommen werden. Nach (2.6) ist aber $L(f) \geqq 0$, und daher nach dem Maximumprinzip $f(P) = f(P_0)$ auf ganz $C_*(P_0)$, d.h. auf $x = x_0$, $z = z_0$, $0 < y < y_2$ einschließlich der Endpunkte, was aber im Widerspruch steht zu $|f| \leqq M + \frac{\alpha}{2}$ für $y \geqq y_2$.

---

* *Zusatz bei der Korrektur:* Für den Sonderfall, daß u, v, w, U, W alle von z unabhängig sind, findet sich bei Nickel (ZAMM **36**, 303–304 (1956)) die dort nicht näher begründete Bemerkung, daß Übergeschwindigkeiten nicht von selbst entstehen können, wenn die Außenströmung drehungsfrei ist.

(b) Würde der Wert $f=M$ im Punkte $P_0 \in G+R'$ angenommen, dann wäre auch $f=M$ auf ganz $C(P_0)$, d.h. auf der Halbgeraden $x=x_0$, $z=z_0$, $0<y<\infty$ einschließlich des Endpunktes $y=0$, was der Voraussetzung $f<M$ an der Wand widerspricht.

Anmerkungen. Beschränkt man sich auf ebene, stationäre Grenzschichtströmungen, dann reduziert sich das Gebiet G auf einen unendlichen Halbstreifen:

$$\begin{aligned} G:&\quad x_1 < x < x_2, & 0<y<\infty \\ R':&\quad x = x_2, & 0<y<\infty \\ R:&\quad x = x_1, & 0<y<\infty \quad \text{und} \\ &\quad x_1 \leqq x \leqq x_2, & y=0. \end{aligned}$$

Wir wollen im folgenden den Fall (b) aus Satz 2.3 betrachten, d.h. es sei $f<M$ an der Wand. Dann kann man auch schreiben

$$M = \sup_{0\leqq y<\infty} \big(u^2(x_1, y) - U^2(x_1)\big).$$

Definiert man noch die Größe e als maximale Übergeschwindigkeit im Anfangsprofil durch

$$e = \sup_{0\leqq y<\infty} \big(u(x_1, y) - U(x_1)\big),$$

dann besteht der Zusammenhang $M=e^2+2eU(x_1)$. Man kann daher Teil (b) aus Satz 2.3 auch so formulieren:

Satz. *Es sei $u\geqq 0$ auf dem hinteren Rande R'. Die äußere Randbedingung für u sei in der strengeren Form (2.8) erfüllt, und an der Wand sei $u^2-U^2<e^2+2eU(x_1)$. Dann gilt*

$$u < \sqrt{e^2+2eU(x_1)+U^2(x)} \quad \text{in} \quad G+R'.$$

In dieser Form hat NICKEL die Abschätzung für die Übergeschwindigkeiten angegeben, allerdings nur unter der Voraussetzung $u\geqq 0$ in $G+R+R'$ und $u=0$ an der Wand, und zwar mit dem Zeichen $\leqq$ in der Ungleichung.

Satz 2.1 lautet für den ebenen Fall:

Satz. *Es sei $u\geqq 0$, $U_x\geqq 0$ in $G+R+R'$. Dann ist sogar $u>0$ in $G+R'$.*

Auch hierin ist eine leichte Verschärfung enthalten gegenüber NICKEL, der an der Wand $u=0$ und $v\leqq 0$ voraussetzte.

Eine unmittelbare Folge des Maximumprinzips ist schließlich der Satz über die Schubspannung u_y im ebenen Fall. Man kann ihn so aussprechen:

Satz. (a) *Es sei G ein beliebiges offenes, nicht notwendig beschränktes Gebiet in der x, y-Ebene, in dem u, v als Lösung der 2-dimensionalen Grenzschichtgleichungen definiert ist, und in dem auch u_{yyy}, $u_{xy}=u_{yx}$ stetig existiert. Wenn u_y in einem Punkt $P_0\in G$ sein Maximum oder Minimum annimmt, dann ist u_y konstant auf ganz $C(P_0)$. Die Aussage dehnt sich aus auf $G+R'$, wenn u, v samt den Ableitungen auf $G+R'$ stetig ist und wenn zusätzlich $u\geqq 0$ auf R' ist.*

(b) *Ist u_y auf dem beschränkten Gebiet G einschließlich des Randes $R+R'$ stetig, dann nimmt u_y sowohl sein Maximum als auch sein Minimum auf R an.*

Ist speziell G der schon oben betrachtete unendliche Halbstreifen, und nimmt u_y in $P_0 \in G+R'$ das Maximum oder Minimum an, dann ist u_y konstant auf der Halbgeraden $x=x_0$, $0 \leqq y < \infty$, d.h. das Profil an der Stelle x_0 hat die Gestalt $u(x_0, y) = \alpha y + \beta$ mit $\alpha = 0$, $\beta = U(x_0)$ wegen der äußeren Randbedingung. Ist aber insbesondere an der Wand $u(x, 0) < U(x)$ vorgeschrieben, dann kann das Maximum oder Minimum überhaupt nicht in $G+R'$ angenommen werden.

3. Instationäre Grenzschichten

Wir betrachten die Grenzschichtgleichungen

$$u_t + u\,u_x + v\,u_y + w\,u_z = -\frac{1}{\varrho} p_x + \nu\,u_{yy}$$

$$w_t + u\,w_x + v\,w_y + w\,w_z = -\frac{1}{\varrho} p_z + \nu\,w_{yy}$$

$$u_x + v_y + w_z = 0.$$

Die Außenströmung wird *zeitunabhängig*, d.h. in der Form $U = U(x,z)$, $W = W(x,z)$ vorausgesetzt und außerdem rotationsfrei[3]. Dann gilt für den Druck die Beziehung $\frac{1}{\varrho} p + \frac{1}{2}(U^2 + W^2) = \text{const}$, und man findet mit dem Operator

$$L \equiv \nu \frac{\partial^2}{\partial y^2} - u \frac{\partial}{\partial x} - v \frac{\partial}{\partial y} - w \frac{\partial}{\partial z} - \frac{\partial}{\partial t}$$

für Lösungen u, v, w der Gleichungen ganz analog wie früher

$$L(u) = -(U\,U_x + W\,W_x), \qquad L(w) = -(U\,U_z + W\,W_z)$$

$$L(u^2 + w^2 - U^2 - W^2) = 2\nu(u_y)^2 + 2\nu(w_y)^2 \geqq 0.$$

Die Sätze 2.1—3 übertragen sich daher ohne weiteres auf instationäre Grenzschichtströmungen.

Der Anschaulichkeit halber wollen wir die Verhältnisse am Beispiel der ebenen Strömungen näher untersuchen. Die Gleichungen lauten dann

$$u_t + u\,u_x + v\,u_y = U\,U' + \nu\,u_{yy}, \tag{3.1}$$

$$u_x + v_y = 0 \tag{3.2}$$

und der Operator reduziert sich auf

$$L \equiv \nu \frac{\partial^2}{\partial y^2} - v \frac{\partial}{\partial y} - u \frac{\partial}{\partial x} - \frac{\partial}{\partial t}.$$

Hierin ist wieder y die „elliptische" Variable $(n=1)$ und x, t sind die „parabolischen" Variablen $(m=2)$. Für Lösungen u, v der Gleichungen (3.1) und (3.2) erhält man dann

$$L(u) = -U\,U', \qquad L(u^2 - U^2) = 2\nu(u_y)^2 \geqq 0, \qquad L(u_y) = 0.$$

Es liegt also formal das gleiche Problem vor wie in Abschnitt 2. Man hat lediglich die drei Raumkoordinaten umzudeuten in zwei Raumkoordinaten und eine Zeitkoordinate. Das Gebiet G und seine Ränder R und R' werden wie folgt

[3] *Zeitabhängige* Außenströmungen $U(x, z, t)$, $W(x, z, t)$ bleiben einer späteren Untersuchung vorbehalten.

erklärt:

$$\begin{array}{llll} G: & x_1 < x < x_2, & 0 < y < \infty, & t_1 < t < t_2 \\ R': & x = x_2, & 0 < y < \infty, & t_1 < t \leqq t_2 \quad \text{und} \\ & t = t_2, & 0 < y < \infty, & x_1 < x \leqq x_2 \\ R: & \multicolumn{3}{l}{\text{Die übrigen Randpunkte von } G.} \end{array}$$

Wir betrachten wieder eine Lösung u, v der Gleichungen, die in $G+R+R'$ stetig ist, und deren Ableitungen soweit sie dort auftreten in $G+R'$ stetig existieren.

Satz 3.1. *Es sei $u(x, y, t) \geqq 0$ in $G+R+R'$, $U(x) > 0$ und $U' > 0$ in $x_1 \leqq x \leqq x_2$. Dann ist sogar $u(x, y, t) > 0$ in $G+R'$.*

Das heißt zu jeder Zeit t ($t_1 < t \leqq t_2$) wird der Wert $u=0$ höchstens an der Wand ($y=0$) oder im Einlaufprofil $u(x_1, y, t)$ angenommen.

Für Satz 3.2 und Satz 3.3 wird die äußere Randbedingung wieder in der strengeren Form vorausgesetzt:

(3.3) $$\lim_{y \to \infty} (u - U) = 0 \qquad \textit{gleichmäßig} \text{ in } x \text{ und } t.$$

Außerdem wird gesetzt

(3.4) $$\bar{u} = \sup_{P \in R} u, \qquad M = \sup_{P \in R} (u^2 - U^2).$$

Es wird also die obere Grenze in folgenden Punkten gebildet: An der Wand ($y=0$) zu allen Zeiten t, dann in den Einlaufprofilen $u(x_1, y, t)$ zu allen Zeiten t, ferner zur Zeit $t=t_1$ in allen Punkten x, y, d.h. für $u(x, y, t_1)$. Bei $u^2 - U^2$ ist es analog.

Satz 3.2. (a) *Es sei $u \geqq 0$ auf R'. Die äußere Randbedingung sei in der strengeren Form (3.3) erfüllt. Ferner sei $U > 0$, $U' \leqq 0$ in $x_1 \leqq x \leqq x_2$. Dann ist $u(x, y, t) \leqq \bar{u}$ in $G+R'$.*

(b) *Wenn zusätzlich an der Wand $u < \bar{u}$ ist, dann gilt sogar $u(x, y, t) < \bar{u}$ in $G+R'$.*

Wenn also keine Rückströmung vorliegt und U nirgends anwächst (bzw. kein Druckabfall stattfindet), dann kann u im ganzen Zeitintervall $t_1 \leqq t \leqq t_2$ nicht größer werden als die Konstante $\bar{u}$.

Satz 3.3. (a) *Es sei $u \geqq 0$ auf R'. Die äußere Randbedingung sei in der strengeren Form (3.3) erfüllt. Dann ist $u^2 \leqq U^2 + M$ in $G+R'$.*

(b) *Wenn diese Ungleichung zusätzlich an der Wand mit dem Zeichen $<$ erfüllt ist, dann gilt sogar $u^2 < U^2 + M$ in $G+R'$.*

Die *a priori* Abschätzung bleibt also auch für zeitabhängige Geschwindigkeitskomponenten $u(x, y, t)$ in derselben Weise bestehen. Trägt man sich die „Energieprofile" $u^2 - U^2$ für festes t und festes x in Abhängigkeit von y auf, so besagt der Satz, daß eventuelle Maxima nicht größer sein können als die Konstante M.

Unmittelbar klar ist im Anschluß an $L(u_y) = 0$ auch das Maximum-Minimumprinzip für die Schubspannung $u_y(x, y, t)$. Wir begnügen uns damit, es für folgendes Gebiet zu formulieren:

$$\begin{array}{llll} G: & x_1 < x < x_2, & y_1 < y < y_2, & t_1 < t < t_2 \\ R': & x = x_2, & y_1 < y < y_2, & t_1 < t \leqq t_2 \quad \text{und} \\ & t = t_2, & y_1 < y < y_2, & x_1 < x \leqq x_2 \\ R: & \multicolumn{3}{l}{\text{Die übrigen Randpunkte von } G.} \end{array}$$

Satz 3.4. (a) *Es sei $u(x, y, t)$, $v(x, y, t)$ als Lösung der 2-dimensionalen, instationären Grenzschichtgleichungen im Gebiet G definiert, und es sei dort auch u_{yyy}, $u_{xy}=u_{yx}$, $u_{ty}=u_{yt}$ stetig. Wenn u_y in einem Punkt $P_0 \in G$ sein Maximum oder Minimum annimmt, dann ist u_y konstant auf ganz $C(P_0)$. Die Aussage dehnt sich aus auf $G+R'$, wenn u, v samt den Ableitungen auf $G+R'$ stetig ist und wenn zusätzlich $u \geqq 0$ auf R' ist.*

(b) *Ist u_y auf dem beschränkten Gebiet G einschließlich des Randes $R+R'$ stetig, dann nimmt u_y sowohl sein Maximum als auch sein Minimum auf R an.*

Für den unendlichen Halbstreifen mit $y_1=0$, $y_2=\infty$ folgt wieder, daß bei Vorliegen eines Maximums oder Minimums das Profil von der Gestalt $u(x_0, y, t_0)=\alpha y+\beta$ ist mit $\alpha=0$, $\beta=U(x_0)$ wegen der äußeren Randbedingung, und daß insbesondere ein Maximum oder Minimum überhaupt nicht angenommen werden kann in $G+R'$, wenn an der Wand $u(x, 0, t)<U(x)$ vorgeschrieben ist.

Anhang. *Zum Beweis von Satz* 1.2. Um die Gültigkeit des Satzes für Punkte des Randes R' einzusehen, braucht man nur den Nirenbergschen Gedankengang zu verfolgen und an wenigen Stellen zu modifizieren. Wir beschränken uns darauf, dies im Falle $n=1$, $m=1$ auszuführen. Für beliebiges n und m verläuft es ganz analog.

Es sei also G ein offenes Gebiet in der x, t-Ebene mit einem Randstück R', das den Bedingungen (a) und (b) aus Abschnitt I genügt. Der Operator lautet

$$L=a\frac{\partial^2}{\partial x^2}+b\frac{\partial}{\partial x}+c\frac{\partial}{\partial t} \qquad \text{mit } a>0 \text{ in } G+R',\ c\leqq 0 \text{ auf } R'.$$

Ähnlich wie bei Nirenberg wird zunächst ein Lemma hergeleitet. Hierzu sei P_0 ein Punkt auf R'. Es gibt dann immer ein $r_0>0$, so daß der Halbkreis H_0 und der Bogen B_0

$$H_0:\quad |x-x_0|^2+|t-t_0|^2<r_0^2, \qquad t\leqq t_0$$
$$B_0:\quad |x-x_0|^2+|t-t_0|^2=r_0^2, \qquad t\leqq t_0$$

ganz zu $G+R'$ gehören.

Lemma. Wenn $L(f)\geqq 0$ ist in ganz $G+R'$, und wenn die Funktion f ihr Maximum M nirgends in H_0 wohl aber auf B_0 im Punkt P_1 annimmt, dann ist notwendig $x_1=x_0$.

Beweis. Man denke sich neue Koordinaten eingeführt mit P_0 als Koordinatenursprung. Dann ist zu zeigen, daß $x_1=0$ sein muß. Der Beweis wird indirekt geführt. Es werde also $x_1\neq 0$ angenommen. Zunächst ist klar, daß P_1 kein Punkt aus G sein kann, weil dann nach Satz 1.1 das Maximum auf ganz $C(P_1)$ angenommen wird, d.h. einem Geradenstück, das Punkte mit H_0 gemeinsam hat. Es bleiben also nur die Endpunkte des Bogens B_0 zu diskutieren. Es sei etwa $P_1=(r_0, 0)$. Da $P_1\in R'$, kann man ein $r>0$ so finden, daß H und B mit

$$H:\quad |x-r_0|^2+t^2<r^2, \qquad t\leqq 0$$
$$B:\quad |x-r_0|^2+t^2=r^2, \qquad t\leqq 0$$

ganz zu $G+R'$ gehört. Sorgt man insbesondere für $0<r<r_0$, dann gehört der Punkt $(0, 0)$ nicht zu $H+B$. Der Bogen B zerfällt in zwei Teilbögen: B' (abgeschlossen) liegt in H_0+B_0, und B'' liegt außerhalb davon. Auf B' liegt kein

Maximumpunkt, d.h. dort ist $f<M$ oder $f\leqq M-\eta$ mit einem genügend klein gewählten positiven η. Da überall $f\leqq M$ ist, hat man speziell

$$f\leqq M-\eta \quad \text{auf } B', \qquad f\leqq M \quad \text{auf } B''.$$

Es wird nun die schon von E. HOPF benutzte Funktion

$$h(x,t)=\exp\{-\alpha(x^2+t^2)\}-\exp\{-\alpha r_0^2\} \qquad (\alpha>0)$$

betrachtet. Sie ist positiv in H_0, null auf B_0 und negativ außerhalb H_0+B_0 für $t\leqq 0$. Man findet

$$\exp\{\alpha(x^2+t^2)\}\,L(h)=4\alpha^2 a(x,t)\,x^2-2\alpha\{a(x,t)+b(x,t)\,x+c(x,t)\,t\}.$$

Da $(0,0)$ nicht zu $H+B$ gehört, ist dort immer $4\alpha^2 a(x,t)\,x^2>0$. Dabei kann man α so groß wählen, daß der in α quadratische Term die in α linearen überwiegt, so daß $L(h)>0$ ist in ganz $H+B$. Wählt man nun ein positives ε so klein, daß mit obigem η stets $\varepsilon h(x,t)\leqq\frac{1}{2}\eta$ in $H+B$ ist, dann hat die Funktion $v\equiv f+\varepsilon h$ folgende Eigenschaften:

Es ist $v<M$ auf B', weil dort $f\leqq M-\eta$ ist, und es ist $v<M$ auf B'', weil dort h negativ ist. Beachtet man $L(u)\geqq 0$, $L(h)>0$, so ergibt sich

$$L(v)>0 \quad \text{auf } H+B, \qquad v<M \quad \text{auf } B.$$

Andererseits ist $h(P_1)=0$, $f(P_1)=M$ und daher $v(P_1)=M$. Die Funktion v nimmt daher auf $B+H$ ihr Maximum nicht auf dem Bogen B sondern nur in einem Punkt $P\in H$ an. In P ist dann $v_{xx}\leqq 0$, $v_x=0$, und falls $P\in G$, auch $v_t=0$. Wenn dagegen P auf dem Rande R' liegt, so ist wenigstens $v_t\geqq 0$. In jedem Falle ist aber wegen $c(x,t)\leqq 0$ auf R' insgesamt $L(v)\leqq 0$ in $P\in H$, im Widerspruch zu $L(v)>0$ auf $H+B$.

Es ist klar, daß das Lemma auch für Halbellipsen anstelle von Halbkreisen gilt, wenn die Hauptachsen parallel zur x- bzw. t-Achse sind. Man hat auch wieder nur die Endpunkte des Bogens zu betrachten, die auf R' liegen. Es genügt dann, Halbkreise zu nehmen, die diesen Ellipsenbogen in den Endpunkten von innen berühren.

Zum Beweis des Satzes selbst werde nun angenommen, daß die Funktion f im Punkt $P\in R'$ ihr Maximum M annimmt. Wäre nun Q ein Punkt von $C(P)$, in dem $f<M$ ist, so gäbe es auf der geradlinigen Verbindung von Q und P, die ganz in $C(P)$ verläuft, einen ersten Punkt P_1 mit $f(P_1)=M$, während zwischen Q und P_1 immer $f<M$ ist. Es seien x_1, t_1 die Koordinaten von P_1. Es werden dann nur noch Punkte betrachtet, die der Menge H_1+B_1 angehören:

$$\begin{aligned} H_1:&\quad |x-x_1|^2+|t-t_1|^2<r_1^2, \qquad t\leqq t_1\\ B_1:&\quad |x-x_1|^2+|t-t_1|^2=r_1^2, \qquad t\leqq t_1, \end{aligned}$$

wobei $r_1>0$ so klein gewählt ist, daß H_1+B_1 ganz zu $G+R'$ gehört.

Es sei nun P_0 ein Punkt auf der Verbindungsgeraden zwischen Q und P_1 mit $|x_0-x_1|\leqq\frac{1}{2}r_1$. In P_0 ist $f<M$, desgleichen ist $f<M$ auf einem Segment A parallel zur t-Achse: $x=x_0$, $t_0-\alpha\leqq t\leqq t_0$ $(\alpha>0)$. Man betrachtet dann die Halbellipse $(t\leqq t_0)$ mit A als der einen Halbachse und mit dem Durchmesser B in

x-Richtung: $t=t_0$, $0\leq|x_0-x|\leq\beta$. Für $\beta=0$ reduziert sich die Halbellipse auf das Geradenstück A, und dort ist $f<M$. Für $\beta=|x_0-x_1|$ liegt auf dem Rande der Punkt P_1 mit $f=M$. Beim stetigen Übergang von $\beta=0$ zu $\beta=|x_0-x_1|$ gibt es dann in der Schar der Halbellipsen eine erste mit folgender Eigenschaft: Auf dem Bogen liegt ein Maximumpunkt, im Innern keiner. Das ist aber gerade die Situation des Lemmas. Der Maximumpunkt hat dann notwendig die Koordinate $x=x_0$, d.h. er liegt auf der Halbachse A, im Widerspruch zu $f<M$ auf A.

Literatur

[1] Görtler, H.: Über die Lösung nichtlinearer partieller Differentialgleichungen vom Reibungsschichttypus. ZAMM **30**, 265—267 (1950).

[2] Hopf, E.: Elementare Bemerkungen über die Lösungen partieller Differentialgleichungen zweiter Ordnung vom elliptischen Typus. Sitzungsber. preuß. Akad. Wiss. **19**, 147—152 (1927).

[3] Hopf, E.: A remark on linear elliptic differential equations of second order. Proc. Amer. Math. Soc. **3**, 791—793 (1952).

[4] Miranda, C.: Equazioni alle derivate parziali di tipo ellittico. Berlin-Göttingen-Heidelberg: Springer 1955.

[5] Nagumo, M., & S. Simoda: Note sur l'inégalité différentielle concernant les équations du type parabolique. Proc. Japan Academy **27**, 536—539 (1951).

[6] Nickel, K.: Einige Eigenschaften von Lösungen der Prandtlschen Grenzschichtdifferentialgleichungen. Arch. Rational Mech. Anal. **2**, 1—31 (1958).

[7] Nirenberg, L.: A strong maximum principle for parabolic equations. Comm. Pure Appl. Math. **6**, 167—177 (1953).

[8] Pucci, C.: Proprietà di massimo e minimo delle soluzioni di equazioni a derivate parziali del secondo ordine di tipo ellitico e parabolico. Consiglio nazionale delle ricerche, Pubblicazioni dell'istituto per le applicazzioni del calcolo N. 517, 1—10 (1958).

[9] Westphal, H.: Zur Abschätzung der Lösungen nichtlinearer parabolischer Differentialgleichungen. Math. Z. **51**, 690—695 (1949).

Institut für Angewandte Mathematik und Mechanik der DVL
Freiburg i. Br.

(Eingegangen am 25. März 1960)

Electric Current and Fluid Spin Created by the Passage of a Magnetosonic Wave

R. P. KANWAL & C. TRUESDELL

Since the discovery of magneto-hydrodynamic waves by ALFVÉN[1], much has been written about them, but one point has failed of general remark[2]: *A magneto-hydrodynamic wave causes the fluid particles to spin about axes parallel to the electric current density it carries; the spin is clockwise or counter-clockwise with respect to the current flow according as the wave normal subtends an acute or an obtuse angle upon the magnetic field vector.* This fact may be verified by inspection of the known special solutions. Here we present a general theory of magneto-hydrodynamic waves, developed in such a way as to draw attention to this new means of generating vorticity in fluid motions.

Weak waves may be studied from five approaches: (a) sinusoidal or other special exact solutions, (b) a linearized theory of small motions, (c) the exact theory of lineal motions, (d) the theory of characteristics, (e) the theory of singular surfaces. All but the last of these have been followed in magneto-hydrodynamics[3]; in particular, a brief but exhaustive treatment by the method of characteristics has been included in FRIEDRICHS & KRANZER's excellent study of waves of all kinds[4]. We shall not expect to find results not contained, at least by implication, in theirs. In developing the subject here by means of the theory of singular surfaces we wish not only to remind physicists of the elegant simplicity of proof offered by the method of CHRISTOFFEL and HUGONIOT, but also to throw light upon the results themselves through expressions which have, at every stage, immediate physical interpretation.

---

[1] [1942].

[2] A relation of this kind, for a linearized theory, has been published by CARSTOIU [1960].

[3] (a) ALFVÉN [1942] [1950, §§ 4.3—4.51] WALÉN [1944, §§ 3—4]. In these special solutions the vorticity and electric current density are everywhere parallel. (b) HERLOFSON [1950, 3], VAN DE HULST [1950, 5], CARSTOIU [1960]. (c) SEGRE [1958, 2]. A general survey of approaches (a), (b), and (c) is given by LÜST [1959, 3]. (References on shock waves are not given; we remark, however, that one of us has determined the relation between current and vorticity generated by a shock [1960, 2]).

[4] [1958, 1, §§ 2—5]. *Cf.* also ONG [1959, 2]. The summary of Russian work given [1959, 1] is too concise to be assessed; we have not been able to see all the references cited there. *Cf.* also [1957, 2].

ALFVÉN discovered transverse waves which propagate in incompressible inviscid fluids of negligible electric resistance at speed a, where[5]

$$a^2 = \frac{\mu H_n^2}{\varrho}, \tag{1}$$

H_n being the component of the magnetic field $\boldsymbol{H}$ normal to the wave-front[6], μ the permeability of the fluid, and ϱ the mass density. In a compressible non-magnetic fluid for which $p = f(\varrho, \eta)$, where p is the pressure and η is the specific entropy, if viscosity and heat conductivity are negligible the only possible waves are longitudinal waves travelling at the speed c, where

$$c^2 = \left(\frac{\partial p}{\partial \varrho}\right)_\eta. \tag{2}$$

In fluids which are both compressible and magnetic, it is natural to expect that waves will generally be neither transverse nor longitudinal, and that their speeds of propagation will differ from a and c.

First we recall the known results obtained by the method of singular surfaces. We do not restate the conditions defining the fluid. Writing $\boldsymbol{u}$ for the velocity field, we restrict "singular surface" to mean a surface across which $\boldsymbol{u}$, ϱ, p, and $\boldsymbol{H}$ are continuous, but some discontinuity is suffered by at least one of the partial derivatives of these quantities with respect to the spatial co-ordinates[7] or the time. In particular, shock waves are not considered in this paper. We adopt the hypotheses of regularity usual in the theory of singular surfaces[8].

Hugoniot-Hadamard theorem: *In a compressible non-magnetic fluid, only the following kinds of singular surfaces are possible:*

1. *Material vortex-sheets.*

2. *Longitudinal waves propagating at speed* $\pm c$ *in any direction.*

If we let $[F]$ stand for the jump of F at the singular surface, while $\boldsymbol{n}$ is the unit normal pointing in the sense in which jumps are calculated, then these sonic waves carry arbitrary jumps of the expansion or the pressure gradient, but the two are related by the equation of state. That is, if

$$\varepsilon \equiv \left[\frac{d \log \varrho}{dn}\right], \qquad \varpi \equiv \left[\frac{dp}{dn}\right], \tag{3}$$

then $\varpi = \varrho \varepsilon c^2$, but either ϖ or ε may be arbitrary. "Material" means "bound to the fluid". A "vortex-sheet" is a surface across which the vorticity or spin

---

[5] M.K.S. units are used throughout this paper.

[6] In much work H appears instead of H_n in the definition of ALFVÉN's speed, but, as appears from the system (4), (5) below, (1) is exact, and other forms are only approximate. The matter is obscured in WALÉN's solutions [1944, §§ 3—4], [1950, § 4.51] but may be cleared by noting that as formulated, they are not invariant under addition of an arbitrary constant vector to the disturbance field, and in fact only disturbances not normal to the constant field can be propagated.

[7] "Singular surface" = "wave of first order" in the terminology of LICHTENSTEIN; in that of HADAMARD, used in [1958, 3], "singular surface" = "wave of second order", except that material singularities affecting only second derivatives with respect to the material co-ordinates are omitted from our treatment.

[8] *E.g.*, [1958, 3].

of the fluid particles is discontinuous but the density gradient is continuous. The discontinuity carried by a vortex-sheet is transversal; the first statement in the Hugoniot-Hadamard theorem asserts that transversal discontinuities cannot propagate but must eternally divide one portion of fluid from another.

Corresponding theorem for ALFVÉN'S waves: *In an incompressible magnetic fluid, only the following kinds of singular surfaces are possible:*

1. *Material vortex-current-sheets. These must be everywhere tangent to the lines of induction (lines of magnetic force); the jumps of vorticity and current density are tangent to the surface but otherwise arbitrary.*

2. *Transverse waves carrying jumps of current density and vorticity satisfying the relations*

$$\varrho U[\boldsymbol{w}] = -\mu H_n[\boldsymbol{j}], \tag{4}$$

$$H_n[\boldsymbol{w}] = -\dot{U}[\boldsymbol{j}], \tag{5}$$

$$\varpi = \mu[\boldsymbol{j}] \cdot \boldsymbol{H}_{\vdash}. \tag{6}$$

Here U is the spead of propagation of the wave, and $\boldsymbol{w}$ and $\boldsymbol{j}$ are the vorticity and the electric current density:

$$\boldsymbol{w} \equiv \operatorname{curl} \boldsymbol{u}, \quad \boldsymbol{j} \equiv \operatorname{curl} \boldsymbol{H}; \tag{7}$$

the jumps $[\boldsymbol{w}]$ and $[\boldsymbol{j}]$ must be tangent to the wave-front, and $\boldsymbol{H}$ is split into normal and tangential components:

$$\boldsymbol{H} = H_n \boldsymbol{n} + \boldsymbol{H}_{\top} = (\boldsymbol{H} \cdot \boldsymbol{n})\,\boldsymbol{n} - \boldsymbol{n} \times \boldsymbol{H}_{\vdash}, \quad \boldsymbol{H}_{\vdash} \equiv \boldsymbol{n} \times \boldsymbol{H}. \tag{8}$$

Thus $\boldsymbol{H}_{\vdash}$ is the vector obtained by rotating the tangential component $\boldsymbol{H}_{\top}$ of $\boldsymbol{H}$ through an angle of 90° counter-clockwise about the normal $\boldsymbol{n}$.

From (4) and (5) we see at once that if $H_n = 0$, *i.e.*, if the wave-front is tangent to $\boldsymbol{H}$, then $U = 0$, and we fall back upon Case 1. This kind of singular surface again divides permanently two portions of fluid, but it must be a surface swept out by lines of induction. The jumps of $\boldsymbol{w}$ and $\boldsymbol{j}$ across it are arbitrary, so long as they be tangential. If $H_n \neq 0$, it follows at once from (4) and (5) that $U^2 = a^2$. *Any wave-front not tangent to the magnetic field is propagated at* ALFVÉN'S *speed* $\pm a$. Note that a^2, unlike c^2, depends upon the orientation of the wave front relative to the magnetic field, being greatest for a wave which moves straight down the lines of induction; then $H_{\vdash} = 0$, and hence $\varpi = 0$ also. Formally, Case 1 may be subsumed under Case 2, since $a = 0$ if $H_n = 0$. In both cases, (4) holds, but in the former, it is vacuous, while in the latter, it gives *the connection between vorticity and electric current density* mentioned in the first sentence of this paper: $[\boldsymbol{w}]$ *and* $[\boldsymbol{j}]$ *point in opposite directions if* $U H_n > 0$, *in the same direction if* $U H_n < 0$, *while their magnitudes stand in the ratio* $\mu|H_n|/\varrho|U|$.

While the known results substantiate the foregoing statements, we cannot point to a published proof for them in this generality. Such a proof is so easy, however, that we omit it.

In this paper we show that (4) *is a universal relation, valid for all kinds of magneto-hydrodynamic singular surfaces*, whether the fluid be compressible or incompressible. Thus the connection found between spin and current is a general one. The relations (5) and (6), and consequently also the speed U, will be seen to need modification when compressibility is taken into account.

Kinematical formulae. For any singular surface as defined above we have the identities of HUGONIOT and HADAMARD, which may be written as follows[9]:

$$\begin{aligned} &[\operatorname{grad} \boldsymbol{u}] = \boldsymbol{n}(\boldsymbol{\alpha} + U\,\varepsilon\,\boldsymbol{n}), && [\dot{\boldsymbol{u}}] = -U(\boldsymbol{\alpha} + U\,\varepsilon\,\boldsymbol{n}), \\ &[\operatorname{div} \boldsymbol{u}] = U\,\varepsilon, && [\operatorname{grad} \varrho] = \varrho\,\varepsilon\,\boldsymbol{n}, \\ &[\boldsymbol{w}] = \boldsymbol{n} \times \boldsymbol{\alpha}, && \boldsymbol{\alpha} = -\boldsymbol{n} \times [\boldsymbol{w}], \end{aligned} \tag{9}$$

where $\dot{\boldsymbol{u}}$ is the acceleration, and where ε, the strength of the condensation, is defined by $(3)_1$. The vector $\boldsymbol{\alpha}$ is *tangent to the singular surface*; the last two equations give an immediate interpretation for this transversal vector in terms of the jump of vorticity.

Preliminary reductions. By the above definition of a singular surface, $[\eta] = 0$. Since dissipative mechanisms are neglected, $\dot{\eta} = 0$, where the dot denotes the material derivative. Hence $[\dot{\eta}] = 0$. It follows by HADAMARD's conditions[9] that if $U \neq 0$, then $[\operatorname{grad} \eta] = 0$. Hence

$$\begin{aligned} [\operatorname{grad} p] &= c^2 [\operatorname{grad} \varrho] + \left(\frac{\partial p}{\partial \eta}\right)_{\varrho} [\operatorname{grad} \eta], \\ &= c^2\,\varrho\,\varepsilon\,\boldsymbol{n}. \end{aligned} \tag{10}$$

(While presentations by engineers and physicists usually assume that $\eta = \text{const.}$, such an assumption is unnecessary, as has been known, in principle, since HUGONIOT's day (1885).)

Second, since $\boldsymbol{H}$ is solenoidal, it follows from WEINGARTEN's theorem[10] that $[\operatorname{grad} \boldsymbol{H}]$ is transversal; in other words, *all magneto-hydrodynamic singular surfaces are current-sheets*, and

$$\begin{aligned} &[\operatorname{grad} \boldsymbol{H}] = \boldsymbol{n}\,\boldsymbol{\beta}, && [\dot{\boldsymbol{H}}] = -U\,\boldsymbol{\beta}, \\ &[\boldsymbol{j}] = \boldsymbol{n} \times \boldsymbol{\beta}, && \boldsymbol{\beta} = -\boldsymbol{n} \times [\boldsymbol{j}], \end{aligned} \tag{11}$$

where $\boldsymbol{\beta}$ is *tangent to the singular surface.* The second of these formulae follows from the first by HADAMARD's conditions[10], and the last two interpret the vector $\boldsymbol{\beta}$ in terms of the jump of electric current density.

Main calculation. When resistance is negligible, FARADAY's law asserts that the lines of induction are material[11]:

$$\dot{\boldsymbol{H}} - \boldsymbol{H} \cdot \operatorname{grad} \boldsymbol{u} + \boldsymbol{H} \operatorname{div} \boldsymbol{u} = 0. \tag{12}$$

An equation expressing balance of linear momentum for an inviscid magnetized fluid is[12]

$$\varrho\,\dot{\boldsymbol{u}} + \operatorname{grad} p + \mu\,\boldsymbol{j} \times \boldsymbol{H} + \varrho\,\boldsymbol{f} = 0. \tag{13}$$

---

[9] [1958, 3, Eq. (9.5)].

[10] [1958, 3, § 4].

[11] This form appears in the earliest magneto-hydrodynamic researches of COWLING; its implications are discussed more formally by WALÉN [1946, § 3.4], TRUESDELL [1950, 4], LUNDQUIST [1952, § IV], COWLING [1957, 1, § 1.31]. This analogy between the induction in a magnetized fluid and the vorticity in a non-magnetized fluid is not to be confused with the connection, emphasized in the present paper, between vorticity and electric current density in the same fluid.

[12] ALFVÉN [1950, 1, Eq. 4.2 (5)], ELSASSER [1950, 2, Eq. (41)], and many later authors.

From (12) and (13), using (11), we see at once that

$$\begin{aligned} &- U\boldsymbol{\beta} - H_n(\boldsymbol{\alpha} + U\varepsilon\boldsymbol{n}) + (H_n\boldsymbol{n} + \boldsymbol{H}_{\top})\,U\varepsilon = 0, \\ &- \varrho\, U(\boldsymbol{\alpha} + U\varepsilon\boldsymbol{n}) + \varrho\, c^2\varepsilon\boldsymbol{n} - \mu(\boldsymbol{n}\times\boldsymbol{\beta})\times\boldsymbol{H} = 0, \end{aligned} \tag{14}$$

where we have assumed the body force $\boldsymbol{f}$ to be continuous. Since $-(\boldsymbol{n}\times\boldsymbol{\beta})\times\boldsymbol{H} = (\boldsymbol{\beta}\cdot\boldsymbol{H}_{\top})\boldsymbol{n} - H_n\boldsymbol{\beta}$, $(14)_1$ and the tangential component of $(14)_2$ become

$$\begin{aligned} &- H_n\boldsymbol{\alpha} - U\boldsymbol{\beta} + U\varepsilon\boldsymbol{H}_{\top} = 0, \\ &- \varrho\, U\boldsymbol{\alpha} - \mu H_n\boldsymbol{\beta} = 0. \end{aligned} \tag{15}$$

It is easier to interpret these equations if we rotate all vectors through a right angle. To the cross products of (15) by $\boldsymbol{n}$, simplified by use of $(8)_3$, $(9)_5$, and $(11)_3$, we subjoin the normal component of $(14)_2$, thus obtaining the ***definitive system:***

$$\varrho\, U[\boldsymbol{w}] + \mu H_n[\boldsymbol{j}] = 0, \tag{16}$$

$$H_n[\boldsymbol{w}] + U[\boldsymbol{j}] = -\varepsilon\, U\boldsymbol{H}_{\vdash}, \tag{17}$$

$$\varrho\,\varepsilon(U^2 - c^2) = -\mu[\boldsymbol{j}]\cdot\boldsymbol{H}_{\vdash}, \tag{18}$$

valid for all singular surfaces. Every quantity occurring has an immediate physical interpretation.

(16) is the same as (4); *it is the universal relation between spin and electric current density* announced above. The interpretation we have given for it in the case of an incompressible fluid was so worded as to remain valid for compressible fluids also. (17) and (18) show how (5) and (6) are modified by the effects of compressibility[13]. The variety of possible singular surfaces is greater.

Classification of singular surfaces. Nothing has been cancelled in deriving (16), (17), and (18), which are all the conditions holding. As we shall see, the possible wave speeds U are determined by them. We assume that $\varrho > 0$, $c^2 > 0$, $H \neq 0$, and we shall tacitly discard cases when the only solutions are $\varepsilon = 0$, $[\boldsymbol{w}] = 0$, $[\boldsymbol{j}] = 0$, *i.e.*, when all quantities are actually continuous.

Case 1. $\varepsilon = 0$: The surface is a vortex-sheet. Then by (18), $[\boldsymbol{j}]$ is parallel to $\boldsymbol{H}_{\top}$ if $H_{\top} \neq 0$.

Case 1a. If $H_n \neq 0$, $\varepsilon = 0$, then $[\boldsymbol{w}]$ is also parallel to $\boldsymbol{H}_{\top}$ if $H_{\top} \neq 0$, and $U^2 = a^2$. If $H_{\top} = 0$, $[\boldsymbol{j}]$ may be arbitrary, but again $U^2 = a^2$. Thus there are *propagating vortex-sheets, purely transversal waves which propagate at* ALFVÉN'S *speed, carrying jumps of vorticity and current density connected by* (4). *If the wave is moving straight down the lines of induction, the jump in current density may be arbitrary both in magnitude and in direction; if the wave is inclined to the magnetic field but not normal to it, the magnitude of the jump of current density may be arbitrary, but its direction must be that of* $\boldsymbol{n}\times\boldsymbol{H}$.

Case 1b. If $H_n = 0$, $\varepsilon = 0$, the system (16)—(18) becomes $\varrho\, U[\boldsymbol{w}] = 0$, $U[\boldsymbol{j}] = 0$, $[\boldsymbol{j}]\cdot\boldsymbol{H}_{\vdash} = 0$. Therefore, since $H_{\vdash} \neq 0$, $[\boldsymbol{j}]$ is parallel to $\boldsymbol{H}$, but $U = 0$, and $[\boldsymbol{w}]$ is arbitrary. Thus *a vortex-sheet tangent to the lines of induction is necessarily*

[13] We may write (18) in the form

$$\varpi = \varrho\varepsilon U^2 + \mu[\boldsymbol{j}]\cdot\boldsymbol{H}_{\vdash},$$

which may be reconciled formally with (6) if we suppose $\varepsilon U^2 \to 0$ as the effects of compressibility become negligible.

material; it can carry any jump of vorticity; the jump of current density must be parallel to $\boldsymbol{H}$ *but is arbitrary in magnitude.* Since $a=0$ when $H_n=0$, this case may be regarded as a limiting case of the former one. (The greater variety of material vortex-sheets possible if the fluid is incompressible follows from the fact that ϖ is then restricted only by (6), while in a compressible fluid $\varpi=\varrho\varepsilon c^2$, so that $\varpi=0$ if $\varepsilon=0$.)

Case 2. $\varepsilon \neq 0$: The surface is a sonic disturbance. If $H_{\vdash}=0$, from (18) we now get $U^2=c^2$, while from (16) and (17) we see that either $U=0$ or $U^2=a^2$. This is a contradiction unless $a^2=c^2$. Hence follows a major difference between the compressible and incompressible cases: *No sonic wave can propagate straight down the lines of induction* except when $a^2=c^2$, that is,

$$\left(\frac{\partial p}{\partial \varrho}\right)_\eta = \frac{\mu H^2}{\varrho}. \tag{19}$$

For an ideal gas, this relation reduces to $\gamma p=\mu H^2$. Such a delicate balance between the static and magnetic pressures would be extraordinary indeed, and henceforth we exclude it.

We take the scalar product of (16) by $H_n \boldsymbol{H}_{\vdash}$, of (17) by $\varrho U \boldsymbol{H}_{\vdash}$, subtract, and eliminate $[\boldsymbol{j}] \cdot \boldsymbol{H}_{\vdash}$ by (18). Thus we obtain an equation determining the speeds of propagation:

$$(U^2-a^2)(U^2-c^2) = U^2 \frac{\mu H_{\mathrm{T}}^2}{\varrho}, \tag{20}$$

where a factor ε has been cancelled.

Case 2a. $U=0$, $\varepsilon \neq 0$: The singularity is a material condensation. Then from (20) we see that $a^2=0$; therefore $H_n=0$. The system (16)—(18) reduces to one member:

$$\varrho\varepsilon c^2 = \mu[\boldsymbol{j}]\cdot\boldsymbol{H}. \tag{21}$$

This case includes Case 1b in the limit when $\varepsilon \to 0$, and it includes the limit case of Case 1a obtained when $a\to 0$. For arbitrary ε, summary of Cases 1b and 2a shows that *every material singular surface is tangent to the lines of induction; the most general such surface may be decomposed into two:*

(i) *A vortex-current-sheet carrying arbitrary jumps in* $\boldsymbol{w}$ *and* $\boldsymbol{j}$.

(ii) *A surface carrying a longitudinal condensation* ε *determined by* (21).

That a jump of current density which is inclined to the magnetic field gives rise to a material current-sheet which is the seat of a condensation is perhaps the most remarkable effect of the interaction between compressibility and magnetism. If the jump of current density lies in the half-plane to the left of $\boldsymbol{H}$, the condensation is positive; if in the right half-plane, negative. Thus a suitable tangential flow of current can render a sheet of induction a material boundary permanently separating a region of dense gas from a region of rare gas. This seems to be the principle which makes possible a magnetic wall.

Case 2b. $H_n=0$, $U\neq 0$, $\varepsilon\neq 0$: The surface is *a propagating condensation always tangent to the lines of induction.* Since $a=0$, (20) yields

$$U^2 = c^2 + \frac{\mu H_{\mathrm{T}}^2}{\varrho} = c^2 + \frac{\mu H^2}{\varrho}, \tag{22}$$

and the system (16)—(18) reduces to the form

$$[\boldsymbol{w}]=0, \qquad [\boldsymbol{j}] = -\varepsilon \boldsymbol{H}_{\vdash}, \qquad \varepsilon H^2 = -[\boldsymbol{j}]\cdot\boldsymbol{H}_{\vdash}, \tag{23}$$

so that *these waves are irrotational; they carry a jump of current density which is of magnitude* $|\varepsilon H|$ *and is in the direction opposite to* $\varepsilon\, \boldsymbol{n}\times\boldsymbol{H}$; *they travel at supersonic speed.* As $H\to 0$, these waves pass continuously into ordinary waves of sound.

The cases considered so far are special and in some way degenerate. We come now to the general case, when the universal relation (4) assumes its full power.

Case 2c. $H_n \neq 0$, $U\neq 0$, $\varepsilon\neq 0$. Two pairs of speeds of propagation are determined by (20), an equation the nature of whose roots has been determined by Friedrichs & Kranzer. Without using their results we see from (16) that now neither $[\boldsymbol{j}]$ nor $[\boldsymbol{w}]$ may vanish, and that these vectors are parallel to each other; by (17), each is parallel to $\boldsymbol{H}_{\vdash}$. The general solution of the system (16) to (18) may be written in terms of a scalar parameter δ:

$$\begin{aligned}
[\boldsymbol{j}] &= \delta(\boldsymbol{n}\times\boldsymbol{H}),\\
[\boldsymbol{w}] &= -\frac{\mu H_n\,\delta}{\varrho\, U}(\boldsymbol{n}\times\boldsymbol{H}),\\
\varepsilon &= \left(\frac{a^2}{U^2}-1\right)\delta.
\end{aligned}\tag{24}$$

Let us think of the condensation ε and the magnetic field $\boldsymbol{H}$ as given. $\boldsymbol{H}$ and the wave-normal $\boldsymbol{n}$ determine two pairs of roots U^2_{fast} and U^2_{slow} of (20); as shown by Friedrichs & Kranzer [14], $U^2_{\text{fast}}>a^2$ and $U^2_{\text{slow}}<a^2$; also $U^2_{\text{fast}}>c^2$ and $U^2_{\text{slow}}<c^2$. By $(24)_3$, then, δ has the same sign as ε for the slow waves but the opposite sign for the fast waves. By (24) we see that for given values of H_n, $\boldsymbol{n}\times\boldsymbol{H}$, and ε there are exactly four possible waves. Summary of the implications of (24) yields the ***main theorem:*** Aside from the special circumstances in which (17) holds, *no wave carrying a condensation can propagate down the lines of induction, but any other direction of propagation is possible.* In the general case, when the wave-front is not tangent to the lines of induction, *there are four possible waves, determined uniquely by* $\boldsymbol{H}$, $\boldsymbol{n}$, *and* ε. *One pair travels at supersonic speed; the other, at subsonic. The jumps of current density and vorticity are parallel to* $\boldsymbol{n}\times\boldsymbol{H}$. *The sense of* $[\boldsymbol{j}]$ *is the same as that of* $\boldsymbol{n}\times\boldsymbol{H}$ *for a slow positive or fast negative condensation, the opposite in the opposite cases. Both members of one pair carry the same condensation and the same jump of current density; they differ only in the sign of* U, *which is either* $+$ *or* $-$ *at will, and in the sense of* $[\boldsymbol{w}]$, *which is opposite to that of* $[\boldsymbol{j}]$ *or the same, according as* $U\boldsymbol{n}\cdot\boldsymbol{H}>0$ *or* $U\boldsymbol{n}\cdot\boldsymbol{H}<0$. *In every case, the magnitudes of all jumps stand in the ratios*

$$|[\boldsymbol{j}]| : |[\boldsymbol{w}]| : |\varepsilon| = |H_{\top}| : \frac{\mu\,|H_n H_{\top}|}{\varrho\,|U|} : \left|\frac{\mu H_n^2}{\varrho\, U^2}-1\right|. \tag{25}$$

[14] Friedrichs & Kranzer have "$\geqq$" and "$\leqq$", but the signs of equality are not possible in the present case. $U^2=a^2\neq 0$ yields $H_{\top}^2=0$ by (20); hence by (18), since $\varepsilon\neq 0$, it follows that $U^2=c^2$. Likewise, $U^2=c^2$ implies that $U^2=a^2$. Thus $c^2=a^2$, but this is the case, excluded here, when (17) holds.

For completeness, we add a proof of the results stated in the text. The case when $a^2=c^2$ is excepted for the reason just given, so that $H_{\top}^2>0$; also $a^2>0$ and $c^2>0$. It is then immediate from (20) that U^2-a^2 and U^2-c^2 have the same sign and do not vanish. Since $U^2_{\text{fast}}\,U^2_{\text{slow}}=a^2c^2$, from $U^2_{\text{fast}}>a^2$ it follows that $U^2_{\text{slow}}<c^2$. Q.E.D.

As $H \to 0$, the fast waves go over into ordinary waves of sound, longitudinal waves carrying an arbitrary condensation in an arbitrary direction and propagating at speed $\pm c$, while the slow waves become indeterminate material singularities. As $\varepsilon \to 0$, both kinds of waves go over into ALFVÉN's waves.

References

1942 ALFVÉN, H.: On the existence of electromagnetic-hydrodynamic waves. Arkiv för Mat. Astron. Fys. B **29**, No. 2, 7 pp.

1944 WALÉN, C.: On the theory of sunspots. Arkiv för Mat. Astron. Fys. A **30**, No. 15, 87 pp.

1946 WALÉN, C.: On the distribution of the solar general magnetic field and remarks concerning the geomagnetism and the solar rotation. Arkiv för Mat. Astron. Fys. A **33**, No. 18, 63 pp.

1950, 1. ALFVÉN, H.: Cosmical Electrodynamics. Oxford.

2. ELASSER, H.: The earth's interior and geomagnetism. Rev. Modern Phys. **22**, 1—35.

3. HERLOFSON, N.: Magneto-hydrodynamic waves in a compressible fluid conductor. Nature **165**, 1020—1021.

4. TRUESDELL, C.: The effect of the compressibility of the earth on its magnetic field. Phys. Rev. (2) **78**, 823.

5. VAN DE HULST, H. C.: Interstellar polarization and magnetohydrodynamic waves. Ch. 6 of Problems of Cosmical Aerodynamics, Proceedings of the Symposium on the Motion of Gaseous Masses of Cosmical Dimension held at Paris, August 16—19, 1949. (Central Air Documents Office, Dayton, Ohio, undated.)

1952 LUNDQUIST, S.: Studies in magneto-hydrodynamics. Arkiv för Fys. **5**, 297—347.

1957, 1. COWLING, T. G.: Magnetohydrodynamics. New York: Interscience.

2. SYROVATSKII, S. I.: Magneto-hydrodynamics [in Russian]. Usp. Fiz. Nauk **62**, 247—303.

1958, 1. FRIEDRICHS, K. O., & H. KRANZER: Notes on magneto-hydrodynamics, VIII. Nonlinear wave motion. AEC R. and D. Report NY 0—6486, New York University.

2. SEGRE, S.: On the formation of magnetohydrodynamic shock waves. Nuovo Cimento (10) **9**, 1054—1057.

3. TRUESDELL, C.: Kinematics of singular surfaces and waves. Univ. Wisconsin MRC Rep. 43. To appear as Chapter C of The Classical Field Theories, Handbuch der Physik, vol. III/1. Berlin-Göttingen-Heidelberg: Springer. In press.

1959, 1. AKHIESER, I. A., & R. V. POLOVIN: Theory of relativistic magnetohydrodynamic waves [in Russian]. Zh. Eksp. Teor. Fiz. **36**, 1845—1852. English translation, JETP **9**, 1316—1320.

2. ONG, R. S.: Characteristic manifolds in three-dimensional unsteady magnetohydrodynamics. Phys. of Fluids **2**, 247—251.

3. LÜST, R. S.: Über die Ausbreitung von Wellen in einem Plasma, Habilitationsschrift München.

1960, 1. CARSTOIU, J.: Hydromagnetic waves in a compressible conductor. Proc. National Acad. Sci. **46**, 131—136.

2. KANWAL, R. P.: Flow behind shock waves in ionized gases. Proc. Roy. Soc. London (in press).

Pennsylvania State University
University Park, Pennsylvania
and
Heat Division, National Bureau of Standards
Washington, D. C.

(Received March 12, 1960)

Stress Tensors in Elastic Dielectrics

R. A. Toupin

1. Introduction

In a previous paper [*1*] we presented a theory of the finite deformation of an elastic, electrically polarizable material. As originally presented, the theory divides a symmetric total stress into two asymmetric components which we called the local stress and the Maxwell stress. Such a decomposition is not necessary but is made for the purpose of introducing a constitutive relation between the local stress and the deformation and polarization of the material. It is the purpose of this note to clarify the theory presented in [*1*] and to exhibit a number of equivalent formulations of its basic equations. In these equivalent formulations, a variety of "stress" tensors appear, and we shall attempt to make clear the relations between them and the stress tensors introduced in [*1*].

The problem of elastic dielectrics affords a simple model of a classical field theory based on a variational principle. Results on this model serve as a guide to the more complex and general dynamical theory of the electromagnetic field in a moving and deforming material which must contain the present statical theory as a special case. Minkowski [*2*] and Abraham [*3*] many years ago proposed different expressions for the electromagnetic stress, energy, momentum, and energy flux in material media. We do not cite the extensive literature which supports one or the other of these expressions. As emphasized by Møller [*5*] and Gyorgyi [*4*], in a material medium, the electromagnetic field forms only part of the physical system. Any division of energy, momentum, stress, and energy flux into electromagnetic and mechanical components is bound to be somewhat arbitrary, and it is fruitless to attempt an independent theory of either component. We trust that the considerations to follow will illustrate these ideas concretely in a simple case where the physical principles are easy to grasp and intuition is trustworthy.

2. Some preliminary formalism

a) The divergence theorem. If $\mathfrak{f}^i$ is a contravariant vector density, *i.e.*, a relative tensor of weight 1, then its *natural divergence* is a scalar density*.

$$\operatorname{div}\mathfrak{f} = \partial_i \mathfrak{f}^i, \qquad \partial_i \equiv \frac{\partial}{\partial x^i}\,. \tag{2.1}$$

* The natural divergence is defined independently of any definition of covariant differentiation. In metric or affinely connected spaces with symmetric affine connection, $\nabla_i \mathfrak{f}^i = \partial_i \mathfrak{f}^i$ provided $\mathfrak{f}$ is a density. *Cf.* Schouten [*6*, Ch. II].

If $\mathscr{s}$ is a surface enclosing a region $\mathscr{v}$ where $\mathfrak{f}$ is continuously differentiable, then

$$\oint \mathfrak{f}^i \, d\mathscr{s}_i = \int \operatorname{div} \mathfrak{f} \, d\mathscr{v}, \tag{2.2}$$

which is the divergence theorem [6, p. 97].

Suppose $\mathscr{v}$ is a region divided by a surface $\mathscr{s}$ across which $\mathfrak{f}$ has a discontinuity $[\mathfrak{f}] = \mathfrak{f}^+ - \mathfrak{f}^-$ where $\mathfrak{f}^+$ and $\mathfrak{f}^-$ are the limiting values of $\mathfrak{f}$ on $\mathscr{s}$. Then if $\mathfrak{f}$ satisfies the smoothness conditions of the divergence theorem in the two parts of $\mathscr{v}$, and if the limits $\mathfrak{f}^+$ and $\mathfrak{f}^-$ exist, we have

$$\oint_{\mathscr{s}} \mathfrak{f}^i \, d\mathscr{s}_i = \int_{v} \operatorname{div} \mathfrak{f} \, d\mathscr{v} + \int_{\mathscr{s} \cap \mathscr{v}} [\mathfrak{f}^i] \, d\mathscr{s}_i, \tag{2.3}$$

where $d\mathscr{s}_i$ points into $\mathscr{v}^+$.

b) Motions. Let

$$\begin{gathered} x^i = x^i(X^A, \lambda), \qquad X^A = X^A(x^i, \lambda) \\ (\boldsymbol{x}/\boldsymbol{X}) \equiv \det \left\| \frac{\partial x^i}{\partial X^A} \right\| \neq 0 \end{gathered} \tag{2.4}$$

be a one parameter family of non-singular continuously differentiable mappings (point transformations) where the indices $i, j, \ldots A, B, \ldots$ have the same range $1, 2, \ldots, n$. In the applications, the variables X^A will be *material coordinates*, the x^i will be spatial coordinates, and n will equal three. We may think of λ as the *time* though this is not a necessary nor even desirable interpretation in purely statical considerations. It is preferable to regard λ simply as an index set which labels the configurations of a given set of material points $\{\boldsymbol{X}\}$. In these preliminaries, we need not assume that the dimension of the space is 3 nor that the space with coordinates x^i is metric in the sense of Riemannian geometry. This will allow application of the general formulae to be derived to problems other than the particular problem of elastic dielectrics considered in §4.

We set

$$\begin{aligned} v^i &\equiv \frac{\partial x^i}{\partial \lambda}, & \grave{X}^A &\equiv \frac{\partial X^A}{\partial \lambda}, \\ x^i_A &\equiv \frac{\partial x^i}{\partial X^A}, & X^A_i &\equiv \frac{\partial X^A}{\partial x^i} \end{aligned} \tag{2.5}$$

and record the identities*

$$\begin{aligned} v^i &= - x^i_A X^A, & \grave{X}^A &= - X^A_i v^i, \\ x^i_A X^A_j &= \delta^i_j, & x^i_A X^B_i &= \delta^A_B, \\ \frac{\partial(\boldsymbol{x}/\boldsymbol{X})}{\partial x^i_A} &= (\boldsymbol{x}/\boldsymbol{X}) X^A_i, & \frac{\partial(\boldsymbol{x}/\boldsymbol{X})}{\partial X^A_i} &= -(\boldsymbol{x}/\boldsymbol{X}) x^i_A, \\ \frac{\partial X^A_i}{\partial x^j_B} &= - X^A_j X^B_i, & \frac{\partial x^i_A}{\partial X^B_j} &= - x^i_B x^j_A. \end{aligned} \tag{2.6}$$

c) Invariant derivatives. Let $f^{ij\cdots}_{k\cdots}(\boldsymbol{x}, \lambda)$ denote a tensor field of arbitrary rank and weight whose components may be functions of the parameter λ as

* *Cf.* TRUESDELL [7].

well as the coordinates $\boldsymbol{x}$. Let

$$\grave{f}^{ij\ldots}_{k\ldots} = \frac{\partial f^{ij\ldots}_{k\ldots}}{\partial \lambda}. \tag{2.7}$$

The *Lie derivative* of $\boldsymbol{f}$ with respect to v^i is a tensor field of the same type as $\boldsymbol{f}$ with components given by*

$$\underset{v}{\pounds} f^{ij\ldots}_{k\ldots} = v^m \partial_m f^{ij\ldots}_{k\ldots} - \partial_m v^i f^{mj\ldots}_{k\ldots} - \partial_m v^j f^{im\ldots}_{k\ldots} - \cdots + \partial_k v^m f^{ij\ldots}_{m\ldots} + \cdots w\, \partial_m v^m f^{ij\ldots}_{k\ldots}, \tag{2.8}$$

where w is the weight of $\boldsymbol{f}$.

Any set of quantities f^Ω, $\Omega = 1, 2, \ldots, N$ which transform by a linear law with coefficients determined by the transformation and its derivatives is called a *geometric quantity* [6, Ch. II, §3]. The definition of the Lie derivative may be extended to arbitrary geometric quantities and has the general form**

$$\underset{v}{\pounds} f^\Omega = v^m \partial_m f^\Omega - F^{\Omega m}_{\Delta n} \partial_m v^n f^\Delta, \tag{2.9}$$

where the $F^{\Omega m}_{\Delta n}$ are constants determined by the law of transformation of the f^Ω. Formula (2.8) is a special case of (2.9). For example, when $\boldsymbol{f}$ is a contravariant vector density, $F^{i n}_{j m} = \delta^n_j \delta^i_m - \delta^i_j \delta^n_m$. The notation of (2.9) is that used by BERGMANN [8].

In a space with symmetric affine connection Γ^i_{jk} and when f^Ω is a set of *tensor* components, we have

$$\nabla_i f^\Omega = \partial_i f^\Omega + F^{\Omega n}_{\Delta m} \Gamma^m_{in} f^\Delta, \tag{2.10}$$

where the $F^{\Omega m}_{\Delta n}$ will have the same values as in (2.9) for corresponding tensors or sets of tensor components f^Ω.

In affinely connected spaces we define the *material derivative* of $\boldsymbol{f}$ by

$$\dot{f}^\Omega \equiv \grave{f}^\Omega + v^m \nabla_m f^\Omega. \tag{2.11}$$

The *convective λ derivative* of $\boldsymbol{f}$ is defined by

$$\overset{*}{f}{}^\Omega \equiv \grave{f}^\Omega + \underset{v}{\pounds} f^\Omega. \tag{2.12}$$

In a space with symmetric affine connection, the partial derivatives of v^i and f^Ω which occur in the definition (2.8) or (2.9) of the Lie derivative may be replaced by covariant derivatives [6, p. 152],

$$\underset{v}{\pounds} f^\Omega = v^m \nabla_m f^\Omega - F^{\Omega n}_{\Delta m} \nabla_n v^m f^\Delta, \tag{2.13}$$

provided, of course, that f^Ω is a tensor field or set of tensor fields for which the covariant derivative is defined. We then have the identity

$$\overset{*}{f}{}^\Omega \equiv \dot{f}^\Omega - F^{\Omega n}_{\Delta m} \nabla_n v^m f^\Delta. \tag{2.14}$$

A geometric quantity f^Ω together with its first partial derivatives $\partial_i f^\Omega$ is a geometric quantity $(f^\Omega, \partial_i f^\Omega)$. In the sense explained by SCHOUTEN [6, p. 105],

* [6, Ch. II, § 10].
** *Cf.* YANO [10].

Lie differentiation and ordinary partial differentiation are commutative:

$$\underset{v}{\pounds}\, \partial_i f^\Omega = \partial_i \underset{v}{\pounds}\, f^\Omega. \tag{2.15}$$

Thus, so also do the operators ∂_i and (*) commute:

$$\overline{\partial_i f^\Omega}^{*} = \partial_i \overset{*}{f}{}^\Omega. \tag{2.16}$$

However, the Lie derivative of an affine connection is given by

$$\underset{v}{\pounds}\, \Gamma^m_{in} = \nabla_i \nabla_n v^m + R_{k\,i\,n}{}^{m} v^k, \tag{2.17}$$

where $\boldsymbol{R}$ is the curvature tensor based on $\boldsymbol{\Gamma}$ and we assume $\Gamma^i_{[jk]}=0$. Taking the Lie derivative of equation (2.10) and using (2.15), we get

$$\underset{v}{\pounds}\, \nabla_i f^\Omega - \nabla_i \underset{v}{\pounds}\, f^\Omega = F^{\Omega\, n}_{\Delta\, m} \underset{v}{\pounds}\, \Gamma^m_{in}, \tag{2.18}$$

from which it is apparent that, even in flat spaces where $R_{k\,i\,n}{}^{m}=0$, Lie differentiation and covariant differentiation do not commute. It follows that covariant differentiation and convected λ differentiation do not commute:

$$\overline{\nabla_k f^\Omega}^{*} - \nabla_k \overset{*}{f}{}^\Omega = F^{\Omega\, n}_{\Delta\, m} \overset{*}{\Gamma}{}^m_{in} f^\Delta. \tag{2.19}$$

As a consequence of the identities (2.6) and the definitions (2.9) and (2.12) we have the useful identities,

$$\dot{X}^K = \overset{*}{X}{}^K \equiv 0, \quad \overset{*}{X}{}^K_i \equiv 0, \quad \overset{*}{x}{}^i_A \equiv 0, \tag{2.20}$$

where the fields X^K are treated as absolute scalars.

It follows from (2.20) that if we define the scalars $F^{AB\ldots}_{C\ldots}$ by

$$F^{AB\ldots}_{C\ldots} \equiv |(\boldsymbol{x}/\boldsymbol{X})|^w X^A_i X^B_j \ldots x^k_C \ldots f^{ij\ldots}_{k\ldots}, \tag{2.21}$$

then

$$f^{ij\ldots}_{k\ldots} = |(\boldsymbol{x}/\boldsymbol{X})|^{-w} x^A_i x^B_j \ldots X^C_k \ldots \dot{F}^{AB\ldots}_{C\ldots} \tag{2.22}$$

where $\boldsymbol{f}$ is a tensor of weight w, and

$$\overset{*}{f}{}^{ij\ldots}_{k\ldots} = |(\boldsymbol{x}/\boldsymbol{X})|^{-w} x^i_A x^j_B \ldots X^C_k \ldots \dot{F}^{AB\ldots}_{C\ldots}, \tag{2.23}$$

since, for scalars, $\overset{*}{F}=\dot{F}$. If the $F^{AB\ldots}_{C\ldots}$ are expressed as functions of the material coordinates X^A and of λ, then $\dot{F}=\dfrac{\partial F}{\partial \lambda}$. It is easy to see that the formulas (2.21) to (2.23) can be written in the abbreviated form

$$\begin{aligned} &f^\Omega = A^\Omega{}_\Delta F^\Delta, \qquad F^\Delta = \overset{-1}{A}{}^\Delta{}_\Omega f^\Omega, \\ &\overset{*}{f}{}^\Omega = A^\Omega{}_\Delta \dot{F}^\Delta, \end{aligned} \tag{2.24}$$

where the coefficients $A^\Omega{}_\Delta$ are certain rational functions of the $x^i{}_A$. In (2.24), the f^Ω may also stand for the ordered set of components of more than one tensor field.

d) Lagrange identities. The field equations which follow from a variational principle are determined by assigning a scalar density $\mathscr{L}$ as a function of the components $(f^\Omega, \partial_i f^\Omega, \partial_i \partial_j f^\Omega, \ldots)$ of a geometric quantity:

$$\mathscr{L} = \mathscr{L}(f^\Omega, \partial_i f^\Omega, \ldots). \tag{2.25}$$

The condition that $\mathscr{L}$ transform as a scalar density, *i.e.*,

$$\mathscr{L}' = |(\boldsymbol{x}/\boldsymbol{x})|^{-1} \mathscr{L} \tag{2.26}$$

under general transformations of the coordinates $\boldsymbol{x}$ leads to a set of algebraic and differential identities satisfied jointly by $\mathscr{L}$, $\partial\mathscr{L}/\partial f^\Omega$, $\partial\mathscr{L}/\partial(\partial_i f^\Omega), \ldots, f^\Omega$, $\partial_i f^\Omega, \ldots$ as shown by Schouten [*6*, Ch. II, §11]*. We restrict attention here to the case where $\mathscr{L}$ depends only on the f^Ω and their first partial derivatives $\partial_i f^\Omega$. This implies no loss in generality since the geometric quantity $(f^\Omega, \partial_i f^\Omega, \ldots, \partial_{i_1}\partial_{i_2}\ldots\partial_{i_n} f^\Omega)$ can always be regarded as the geometric quantity $(g^{\Omega'}, \partial_i g^{\Omega'})$ where $g^{\Omega'} = (f^\Omega, \partial_i f^\Omega, \ldots, \partial_{i_1}\partial_{i_2}\ldots\partial_{i_{n-1}} f^\Omega)$. However, in the applications, the geometric quantity f^Ω which appears in (2.27) below is assumed to have a linear homogeneous law of transformation corresponding to the special case where f^Ω stands for the ordered set of components of a set of *tensor* fields. Thus we assume that $\mathscr{L}$ has the form

$$\mathscr{L} = \mathscr{L}(f^\Omega, \partial_i f^\Omega). \tag{2.27}$$

We then have

$$\underset{v}{\pounds}\, \mathscr{L} = \partial_k (v^k \mathscr{L}) = \frac{\partial\mathscr{L}}{\partial f^\Omega} \underset{v}{\pounds}\, f^\Omega + \frac{\partial\mathscr{L}}{\partial f_i^\Omega} \underset{v}{\pounds}\, f_i^\Omega \tag{2.28}$$

which must hold for arbitrary v^k. In (2.28) and henceforth we use the notation $f_i^\Omega = \partial_i f^\Omega$. By equating coefficients of v^k, $\partial_m v^k$, and $\partial_m \partial_n v^k$ on both sides of $(2.28)_2$ we get the identities**:

$$\partial_k \mathscr{L} = \frac{\partial\mathscr{L}}{\partial f^\Omega} + \frac{\partial\mathscr{L}}{\partial f_i^\Omega} \partial_k f_i^\Omega, \tag{2.29}$$

$$\delta_j^i \mathscr{L} = \frac{\partial\mathscr{L}}{\partial f^\Omega} f_j^\Omega - F^{\Omega\, i}_{\Delta\, j} \left(\frac{\partial\mathscr{L}}{\partial f^\Omega} f^\Delta + \frac{\partial\mathscr{L}}{\partial f_k^\Omega} f_k^\Delta \right), \tag{2.30}$$

$$0 = \frac{\partial\mathscr{L}}{\partial f^\Omega_{(k}} F^{\Omega\, i)}_{\Delta\, j} f^\Delta. \tag{2.31}$$

In field theories based on a variational principle with Lagrangean $\mathscr{L}$ the set of quantities,

$$_{\mathrm{C}}t^i{}_j \equiv \delta_j^i \mathscr{L} - \frac{\partial\mathscr{L}}{\partial f_i^\Omega} f_j^\Omega, \tag{2.32}$$

is called the *canonical stress matrix*. In dynamical theories where the underlying space of points $\boldsymbol{x}$ includes one dimension of time, $_{\mathrm{C}}t^i{}_j$ is called the canonical *stress-energy-momentum matrix*. Only in special cases will $_{\mathrm{C}}t^i{}_j$ transform as a mixed tensor density of weight 1 under general transformations of the coordinates x^i.

* *Cf.* also Thomas [*9*, Chs. VI, VII].

** The non-trivial identities (2.30) and (2.31) were obtained by Einstein [*13*] for the case when f^Ω is a symmetric, non-singular, second order absolute tensor. The generalization of these identities to an arbitrary geometric quantity f^Ω upon which the remaining identities of this section rest was given by Bergmann [*8*]. Various special cases of these identities are well known in differential geometry. See, *e.g.*, Thomas [*9*] and Schouten [*6*].

On computing the natural divergence of ${}_{\mathrm{C}}\mathfrak{t}^i{}_j$ and using the identity (2.29), we find that

$$\partial_i({}_{\mathrm{C}}\mathfrak{t}^i{}_j) \equiv L_\Omega f^\Omega_j, \tag{2.33}$$

where L_Ω is the *Lagrange derivative* of $\mathscr{L}$ with respect to f^Ω defined by

$$L_\Omega \equiv \frac{\partial \mathscr{L}}{\partial f^\Omega} - \partial_i \left(\frac{\partial \mathscr{L}}{\partial f^\Omega_i} \right). \tag{2.34}$$

The Lagrange derivative of $\mathscr{L}$ with respect to a tensor of weight w is a tensor of weight $1-w$. Integrating the last term in the identity (2.30) by parts, introducing the Lagrange derivative (2.34), and making use of the identity (2.31), we get the identity

$${}_{\mathrm{C}}\mathfrak{t}^i{}_j \equiv - L_\Omega F^{\Omega}_{\Delta}{}^i_j f^\Delta - \partial_k \left(\frac{\partial \mathscr{L}}{\partial f^\Omega_{[k}} F^{\Omega}_{\Delta}{}^{i]}_j f^\Delta \right). \tag{2.35}$$

If one now computes the divergence of ${}_{\mathrm{C}}\mathfrak{t}^i{}_j$, the last term contributes nothing because of its antisymmetry in the indices i and k and one gets using (2.33)

$$L_\Omega f^\Omega_j + \partial_i (L_\Omega F^{\Omega}_{\Delta}{}^i_j f^\Delta) \equiv 0, \tag{2.36}$$

which is BERGMANN'S *identity* [8]. Equivalently, if we put

$$\mathfrak{T}^i{}_j \equiv {}_{\mathrm{C}}\mathfrak{t}^i{}_j + L_\Omega F^{\Omega}_{\Delta}{}^i_j f^\Delta = - \partial_k \left(\frac{\partial \mathscr{L}}{\partial f^\Omega_{[k}} F^{\Omega}_{\Delta}{}^{i]}_j f^\Omega \right), \tag{2.37}$$

we see that (2.36) is equivalent to the identity

$$\partial_i \mathfrak{T}^i{}_j \equiv 0. \tag{2.38}$$

Given any scalar density function of a geometric quantity f^Ω and its derivatives f^Ω_k, formula (2.37) constitutes a general rule for the construction of an array of n^2 quantities whose divergence vanishes *identically* as a consequence of the assumed law of transformation for $\mathscr{L}$. The f^Ω need not be a set of tensor fields but may also include more general geometric quantities such as an affine connection.

Special interest is attached to the case where one of the quantities included in the set f^Ω is a symmetric, non-singular tensor g_{mn} which we may think of as the metric tensor for the underlying space, though this is not a necessary interpretation for the present analysis. In this case, we indicate the functional dependence of $\mathscr{L}$ by writing

$$\mathscr{L} = \mathscr{L}(g_{mn}, \partial_k g_{mn}, \varphi^{\Omega'}, \varphi^{\Omega'}_k) \tag{2.39}$$

and separating the tensor g_{mn} explicitly from the total set f^Ω. If f^Ω is a geometric quantity, the remaining set of components $\varphi^{\Omega'}$ is also a geometric quantity whose Lie derivative we denote by

$$\underset{v}{\pounds}\, \varphi^{\Omega'} = v^k \partial_k \varphi^{\Omega'} - \Phi^{\Omega'}_{\Delta'}{}^i_j \partial_i v^j \varphi^{\Delta'}. \tag{2.40}$$

We introduce the special notation $\mathfrak{t}^{mn}$ for twice the Lagrange derivative of $\mathscr{L}$ with respect to g_{mn}:

$$\mathfrak{t}^{mn} \equiv 2 \left[\frac{\partial \mathscr{L}}{\partial g_{mn}} - \partial_k \left(\frac{\partial \mathscr{L}}{\partial(\partial_k g_{mn})} \right) \right]. \tag{2.41}$$

Bergmann's identity (2.36) applied to a Lagrange function having the form (2.39) yields Trautman's identity [*11*]:

$$\nabla_i \mathrm{t}^i_{\ j} = L_{\Omega'}\,\varphi^{\Omega'}_j + \partial_i (L_{\Omega'}\,\Phi^{\Omega'\ i}_{\Delta'\ j}\,\varphi^{\Delta'}), \tag{2.42}$$

where $\mathrm{t}^i_{\ j} = g_{jk}\,\mathrm{t}^{ik}$, and ∇_i denotes covariant differentiation based on the Christoffel symbols of the tensor g_{mn}.

The Lagrange derivative of $\mathscr{L}$ with respect to an absolute tensor field is a tensor field of weight 1, *i.e.*, a tensor density. For reasons to appear later in the application of these general formulae, we shall call the symmetric tensor density t^{ij} defined in (2.41) and the associated absolute tensor $t^{ij} = \mathrm{t}^{ij}/\sqrt{|\det g_{mn}|}$, the *Cauchy stress*. From (2.42) we see that a sufficient condition for the covariant divergence of the Cauchy stress to vanish is that $L_{\Omega'} = 0$, *i.e.*, that the Lagrange derivative of $\mathscr{L}$ with respect to each of the remaining field components $\varphi^{\Omega'}$ vanish. From (2.33) it is apparent that $L_\Omega = 0$, which, in the case of a Lagrangean having the form (2.39), includes the condition $\mathrm{t}^{mn} = 0$, is a sufficient condition that the ordinary divergence of the canonical stress matrix vanish. We have previously noted that the ordinary divergence of the tensor $\mathfrak{T}^i_{\ j}$ vanishes *identically*.

Henceforth we restrict attention to the case where $\mathscr{L}$ is independent of the partial derivatives of g_{mn}:

$$\mathscr{L} = \mathscr{L}(g_{mn}, \varphi^{\Omega'}, \varphi^{\Omega'}_m). \tag{2.43}$$

When $\mathscr{L}$ has this form we can solve the identity (2.30) for the Cauchy stress.

$$\mathrm{t}^i_{\ j} = {}_{\mathrm{c}}\mathrm{t}^i_{\ j} + \Phi^{\Omega'\ i}_{\Delta'\ j}\left(\frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}}\,\varphi^{\Delta'} + \frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}_k}\,\varphi^{\Delta}_k\right). \tag{2.44}$$

Now, by definition, the Cauchy stress is symmetric.

$$\mathrm{t}^{[ij]} \equiv 0, \tag{2.45}$$

which, by (2.44) implies that

$$_{\mathrm{c}}\mathrm{t}^{[ij]} + \Phi^{\Omega'\ [i}_{\Delta'\ k}\, g^{j]k}\left(\frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}}\,\varphi^{\Delta'} + \frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}_m}\,\varphi^{\Delta'}_m\right) = 0. \tag{2.46}$$

Thus, in general, the canonical stress ${}_{\mathrm{c}}\mathrm{t}^{ij} \equiv g^{jk}\,{}_{\mathrm{c}}\mathrm{t}^i_{\ k}$ is not symmetric. Equation (2.46) is an explicit formula for its antisymmetric part when the Lagrangean has the special form (2.43).

3. Field equations and boundary condition of a variational principle

Let $\mathscr{L}$ have the special form (2.43) and consider the integral

$$I(\lambda) \equiv \int_{v(\lambda)} \mathscr{L}\, dv(\lambda) \tag{3.1}$$

where $\varphi^{\Omega'} = \varphi^{\Omega'}(\boldsymbol{x}, \lambda)$ but the metric components $g_{mn}(\boldsymbol{x})$ are assumed to be independent of λ. The region of integration $v(\lambda)$ in (3.1) corresponds to a fixed

set of material points X^K and thus depends on the parameter λ. Let $v(\lambda)$ be divided by a surface of discontinuity $\mathscr{S}(\lambda)$ where we assume that $\mathscr{S}(\lambda)$ corresponds also to a fixed set of points X^K. The metric will be assumed continuously differentiable throughout $v(\lambda)$, but the fields $\varphi^{\Omega'}$ will be assumed continuously differentiable only in each of the regions v^+ and v^- and may have ordinary discontinuities at $\mathscr{S}(\lambda)$. Under these conditions we have

$$\frac{dI}{d\lambda} = \int \overset{*}{\mathscr{L}}\, dv = \int \left(\frac{\partial \mathscr{L}}{\partial g_{mn}} \overset{*}{g}_{mn} + \frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}} \overset{*}{\varphi}{}^{\Omega'} + \frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}_m} \overset{*}{\varphi}{}^{\Omega'}_m \right) dv. \tag{3.2}$$

Since $\overset{*}{\varphi}{}^{\Omega'}_m = \partial_m \overset{*}{\varphi}{}^{\Omega'}$, and $\overset{*}{g}_{mn} = \partial_m v^k g_{kn} + \partial_n v^k g_{mk}$, we can integrate certain of the terms in (3.2) by parts and apply the divergence theorem (2.3) to get

$$\begin{aligned} \frac{dI}{d\lambda} = \int_v \left(-\nabla_m \mathrm{t}^m{}_k v^k + L_{\Omega'} \overset{*}{\varphi}{}^{\Omega'}\right) dv + \oint_s \left(\mathrm{t}^m{}_k v^k + \frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}_m} \overset{*}{\varphi}{}^{\Omega'}\right) d s_m - \\ - \int_{\mathscr{S} \cap v} \left[\mathrm{t}^m{}_k v^k + \frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}_m} \overset{*}{\varphi}{}^{\Omega'}\right] d\mathscr{S}_m. \end{aligned} \tag{3.3}$$

Let the derivatives (variations) v^k and $\overset{*}{\varphi}{}^{\Omega'}$ be subject to a set of N *constraints*

$$0 = C^\omega = a^\omega{}_k v^k + b^\omega{}_{\Omega'} \overset{*}{\varphi}{}^{\Omega'}, \qquad \omega = 1, 2, \ldots, N. \tag{3.4}$$

By introducing the Lagrange multipliers η_ω in the usual way we obtain the result:

A sufficient condition that the integral $I(\lambda)$ have a stationary value for arbitrary variations $(v^k, \overset{*}{\varphi}{}^{\Omega'})$ consistent with the constraints $C^\omega = 0$, vanishing on the boundary $s(\lambda)$ of $v(\lambda)$, and continuous at $\mathscr{S}(\lambda) = 0$, i.e., $[v^k] = 0$, $[\overset{*}{\varphi}{}^{\Omega'}] = 0$, is that the field equations

$$\begin{aligned} -\nabla_m \mathrm{t}^m{}_k + \eta_\omega a^\omega{}_k &= 0, \\ L_{\Omega'} + \eta_\omega b^\omega{}_{\Omega'} &= 0, \end{aligned} \tag{3.5}$$

be satisfied at each interior point of the regions v^+ and v^- and that, at the surface of discontinuity $\mathscr{S}$, the boundary conditions

$$\begin{aligned} [\mathrm{t}^m{}_k]\, n_m - \eta_\omega a^\omega{}_k &= 0, \\ \left[\frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}_m}\right] n_m + \eta_\omega b^\omega{}_{\Omega'} &= 0 \end{aligned} \tag{3.6}$$

be satisfied at each point of $\mathscr{S}(\lambda)$ interior to $v(\lambda)$.

Remarks. Suppose that the $\varphi^{\Omega'}$ are the components of a set of tensor fields and introduce the scalars $\Phi^{\Omega'} = \overset{-1}{A}{}^{\Omega'}{}_{\Delta'} \varphi^{\Delta'}$ as in (2.24). Then one has $\overset{*}{\varphi}{}^{\Omega'} = A^{\Omega'}{}_{\Delta'} \dot{\Phi}^{\Delta'}$ and from $(2.6)_1$, $v^k = -x^k_A X^A$. Thus the assumption that the variations $(v^k, \overset{*}{\varphi}{}^{\Omega})$ are continuous at $\mathscr{S}(\lambda)$ is seen to be equivalent to the assumption that the variations $(\dot{\Phi}^{\Omega'}, X^A)$ are continuous at $\mathscr{S}(\lambda)$ provided one assumes that $[A^{\Omega'}{}_{\Delta'}] = 0$ which holds when the X^A_i are continuous. A different set of boundary conditions

arises if we write $dI/d\lambda$ in the form given by WEISS [12],

$$
\begin{aligned}
\frac{dI}{d\lambda} &= \int\limits_{v} \left\{ \frac{\partial \mathscr{L}}{\partial \lambda} + \partial_k (v^k \mathscr{L}) \right\} dv \\
&= \int\limits_{v} \left(\frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}} \dot{\varphi}^{\Omega'} + \frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}_k} \dot{\varphi}^{\Omega'}_k \right) dv + \oint\limits_{s} v^k \mathscr{L}\, ds_k - \int\limits_{\mathscr{S} \cap v} [v^k \mathscr{L}]\, d\mathscr{S}_k \\
&= \int\limits_{v} L_{\Omega'} \dot{\varphi}^{\Omega'} dv + \oint\limits_{s} \left\{ \frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}_k} \dot{\varphi}^{\Omega'} + v^k \mathscr{L} \right\} ds_k - \int\limits_{\mathscr{S} \cap v} \left[\frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}_k} \dot{\varphi}^{\Omega'} + v^k \mathscr{L} \right] d\mathscr{S}_k \\
&= \int\limits_{v} L_{\Omega'} \dot{\varphi}^{\Omega'} dv + \oint\limits_{s} \left\{ \frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}_k} \overset{*}{\varphi}^{\Omega'} + \mathfrak{t}^k{}_j v^j \right\} ds_k - \int\limits_{\mathscr{S} \cap v} \left\{ \left[\frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}_k} \overset{*}{\varphi}^{\Omega'} \right] + [\mathfrak{t}^k{}_j v^j] \right\} d\mathscr{S}_k ,
\end{aligned}
\tag{3.7}
$$

where, in the last equation of (3.7), the variation $\overset{*}{\varphi}^{\Omega'} \equiv \dot{\varphi}^{\Omega'} + v^k \partial_k \varphi^{\Omega'}$ is not a tensor unless all the $\varphi^{\Omega'}$ are absolute scalars and therefore, in general, has no particular geometric significance*. One then sees from (3.7) that a sufficient condition that $I(\lambda)$ have a stationary value for all variations $(\overset{*}{\varphi}^{\Omega'}, v^k)$ which vanish on s and are continuous in $v(\lambda)$ is that $L_{\Omega'} = 0$ in v^+ and v^- and $[\mathfrak{t}^i{}_j] n_i = 0$, $\left[\frac{\partial \mathscr{L}}{\partial \varphi^{\Omega'}_k} \right] n_k = 0$ on $\mathscr{S}(\lambda)$, provided there are no constraints on the variations $(\overset{*}{\varphi}^{\Omega'}, v^k)$.

4. Elastic dielectrics

The formalism of §2 and §3 will now be applied to a special case to reproduce the equations of the theory of large elastic deformations of a solid dielectric material considered previously in [1].

Let the space of points $\boldsymbol{x}$ be Euclidean three-dimensional space, and let the Lagrangean function have the form

$$
\mathscr{L} = \mathfrak{W}(X^K, X^K_i, \mathfrak{P}^i, g_{ij}) + \mathfrak{P}^i \partial_i \varphi - (\varepsilon_0 \sqrt{\mathfrak{g}}/2)\, g^{ij} \partial_i \varphi\, \partial_j \varphi \tag{4.1}
$$

where g_{ij} are the components of the Euclidean metric tensor, φ is an absolute scalar field called the *electric potential*, $\mathfrak{P}^i$ is a contravariant vector density with the transformation law,

$$
\mathfrak{P}^{i'} = |(\boldsymbol{x}'/\boldsymbol{x})|^{-1} \frac{\partial x^{i'}}{\partial x^j} \mathfrak{P}^j, \tag{4.2}
$$

called the *polarization density*, and $\mathfrak{W}$ is a scalar density we call the *energy density of elastic deformation and polarization*. Note that $\mathfrak{W}$ is independent of φ and its derivatives, the central term in (4.1) is independent of the *deformation gradients* X^A_i, and the last group of terms in (4.1) depends only on g_{ij} and $\partial_i \varphi$. ε_0 is a constant, and, as usual, $\mathfrak{g}$ denotes the determinant of the metric. Since we assume that the metric is Euclidean, there exist coordinate systems, the rectangular Cartesian frames, for which $g_{ij} = \delta_{ij}$ everywhere. However, it is to be noted that $\overset{*}{\delta}_{ij} \neq 0$ except for rigid variations. In some works on electromagnetic theory it is customary to represent the polarization field by the absolute vector field $P^i = \mathfrak{P}^i/\sqrt{\mathfrak{g}}$. Actually, it is more convenient to work with the vector density $\mathfrak{P}^i$

* WEISS [12, p. 105] states that the $\dot{\varphi}^{\Omega'}$ are "more fundamental" than the variations $\varphi^{\Omega'}$ and gives a reason we do not understand. The variations $\overset{*}{\varphi}^{\Omega'}$, on the other hand, have an invariant significance even in non-metric or non-affinely connected spaces.

which has the transformation law (4.2). If one restricts the coordinate transformations to the unimodular ones, no distinction need be maintained between tensor densities and absolute tensors. The scalar density $\mathfrak{W}$ is related to the absolute scalar function Σ introduced in [1] by

$$\mathfrak{W} = \varrho_0 (\boldsymbol{x}/\boldsymbol{X})^{-1} \Sigma \tag{4.3}$$

where ϱ_0 is the density of mass in the undeformed or natural state of the material.

A material characterized by the field equations and boundary conditions (3.5) and (3.6) with a Lagrange function having the form (4.1) will be called a *perfectly elastic dielectric* provided the constraints do not include the condition $\mathfrak{P}=0$.

The canonical stress in an elastic dielectric is given by

$$\begin{aligned} {}_{\mathrm{C}}t^i{}_j &= \delta^i_j \mathscr{L} - \frac{\partial \mathscr{L}}{\partial \varphi_i} \varphi_j - \frac{\partial \mathscr{L}}{\partial X^K_i} X^K_j \\ &= \left(\delta^i_j \mathfrak{W} - \frac{\partial \mathfrak{W}}{\partial X^K_i} X^K_j\right) + \mathfrak{P}^i E_j + \varepsilon_0 \sqrt{\mathfrak{g}}\, E^i E_j - \delta^i_j \left(\mathfrak{P} \cdot E + \frac{\varepsilon_0}{2} \sqrt{\mathfrak{g}}\, E^2\right), \end{aligned} \tag{4.3}$$

where we have introduced the *electric field*, $E_i \equiv -\partial_i \varphi$, $E^i \equiv g^{ij} E_j$, $\mathfrak{P} \cdot \boldsymbol{E} \equiv \mathfrak{P}^i E_i$, *etc.*

Now in [1] we introduced the absolute scalar field π^i defined by

$$\pi^i \equiv (\boldsymbol{x}/\boldsymbol{X})\, \mathfrak{P}^i/\varrho_0 \tag{4.4}$$

and wrote Σ in the form

$$\Sigma = \hat{\Sigma}(X^K, x^i_A, \pi^i, g_{ij}). \tag{4.5}$$

We then introduced a *local stress tensor* ${}_{\mathrm{L}}t^i{}_j$ and a *local field* ${}_{\mathrm{L}}E_i$ defined by

$$ {}_{\mathrm{L}}t^i{}_j \equiv \varrho \frac{\partial \hat{\Sigma}}{\partial x^j_A} x^i_A, \qquad {}_{\mathrm{L}}E_i = -\frac{\partial \hat{\Sigma}}{\partial \pi^i}, \tag{4.6}$$

where $\varrho = |(\boldsymbol{x}/\boldsymbol{X})|^{-1} \varrho_0$ is the density of mass in the present configuration. From (4.2), (4.4), (4.5), (4.6) and the identities given in § 2 we derive the identities:

$$\begin{aligned} \delta^i_j \mathfrak{W} - \frac{\partial \mathfrak{W}}{\partial X^K_i} X^K_j &\equiv {}_{\mathrm{L}}t^i{}_j - \delta^i_j\, {}_{\mathrm{L}}\boldsymbol{E}\ \mathfrak{P}, \\ \frac{\partial \mathfrak{W}}{\partial \mathfrak{P}^i} &\equiv -\, {}_{\mathrm{L}}E_i. \end{aligned} \tag{4.7}$$

Substituting these relations into (4.3), we get

$$ {}_{\mathrm{C}}t^i{}_j = {}_{\mathrm{L}}t^i{}_j + \mathfrak{P}^i E_j + \varepsilon_0 \sqrt{\mathfrak{g}}\, E^i E_j - \delta^i_j\, (\boldsymbol{E} + {}_{\mathrm{L}}\boldsymbol{E}) \cdot \mathfrak{P} - (\varepsilon_0 \sqrt{\mathfrak{g}}/2)\, \delta^i_j E^2, \tag{4.8}$$

which exhibits an explicit relation between the canonical stress and the local stress.

For elastic dielectrics, the symmetric Cauchy stress as given by (2.44) turns out to have the form

$$\begin{aligned} t^i{}_j &= {}_{\mathrm{C}}t^i{}_j + (\delta^i_m \delta^n_j - \delta^i_j \delta^n_m) \frac{\partial \mathscr{L}}{\partial \mathfrak{P}^n} \mathfrak{P}^m, \\ &= {}_{\mathrm{C}}t^i{}_j - (E_j + {}_{\mathrm{L}}E_j)\, \mathfrak{P}^i + \delta^i_j (\boldsymbol{E} + {}_{\mathrm{L}}\boldsymbol{E}) \cdot \mathfrak{P}. \end{aligned} \tag{4.9}$$

Substituting from (4.8) for the canonical stress in (4.9), we get

$$\mathfrak{t}^i{}_j = {}_{\mathsf{L}}\mathfrak{t}^i{}_j - {}_{\mathsf{L}}E_j\,\mathfrak{P}^i + \varepsilon_0\sqrt{\mathfrak{g}}\,E^i E_j - (\varepsilon_0\sqrt{\mathfrak{g}}/2)\,\delta^i_j\,E^2 = 2\,\frac{\partial\mathscr{L}}{\partial g_{im}}\,g_{jm}. \tag{4.10}$$

The field equations and boundary conditions are obtained from the general formulae of §3 as follows. First note that $\overset{*}{X}{}^K = 0$, $\overset{*}{X}{}^A_i = 0$, so that the constaints (3.4) corresponding to these conditions have a simple form. We then consider a region $\mathscr{v}(\lambda)$ filled with elastic dielectric which might consist of two disimilar materials with the common boundary $\mathscr{S}(\lambda)$. In $\mathscr{v}^+$ and $\mathscr{v}^-$ we assume the variations $\overset{*}{\mathfrak{P}}{}^i$ and $\overset{*}{\varphi}$ are arbitrary. We then get the field equations and boundary conditions:

$$\begin{gathered}\nabla_j\mathfrak{t}^j{}_k = 0,\\ {}_{\mathsf{L}}E_i + E_i = 0,\\ \varepsilon_0\sqrt{\mathfrak{g}}\,\nabla^2\varphi - \operatorname{div}\mathfrak{P} = 0,\end{gathered} \tag{4.11}$$

which must hold at each point of $\mathscr{v}^+$ and $\mathscr{v}^-$, and

$$[\mathfrak{t}^j{}_k]\,n_j = 0, \qquad [\varepsilon_0\sqrt{\mathfrak{g}}\,E^i + \mathfrak{P}^i]\,n_i = 0 \tag{4.12}$$

which must hold at each point of $\mathscr{S}$ contained in $\mathscr{v}$.

When the field equation $(4.11)_2$ is satisfied, the Cauchy stress tensor (4.10) reduces to the form

$$\mathfrak{t}^i{}_j = {}_{\mathsf{L}}\mathfrak{t}^i{}_j + E_j\,\mathfrak{P}^i + (\varepsilon_0\sqrt{\mathfrak{g}}/2)\,\delta^i_j\,E^2 \tag{4.13}$$

given in [*1*]. Thus the field equations and boundary conditions (4.11) and (4.12) are equivalent to those of [*1*]. When the field equation $(4.11)_2$ is satisfied, it follows from $(4.9)_2$ that the Cauchy stress and the canonical stress are equal in elastic dielectrics.

From the symmetry of the Cauchy stress and (4.10) we conclude that

$${}_{\mathsf{L}}\mathfrak{t}^{[ij]} - {}_{\mathsf{L}}E^{[i}\,\mathfrak{P}^{j]} = 0, \tag{4.14}$$

which is equation (10.36) of [*1*]. The derivation of (4.14) given in [*1*] shows how this identity follows from the invariance of Σ under rigid motions.

It is natural to introduce still another stress tensor in elastic dielectrics which we define by

$${}_{\mathsf{M}}\mathfrak{t}^{ij} \equiv {}_{\mathsf{L}}\mathfrak{t}^{ij} - {}_{\mathsf{L}}E^i\,\mathfrak{P}^j. \tag{4.15}$$

Let us call ${}_{\mathsf{M}}\mathfrak{t}^{ij}$, the *material stress*. By (4.14) it is always symmetric. Thus the Cauchy stress can be given the alternative decomposition

$$\mathfrak{t}^{ij} = {}_{\mathsf{M}}\mathfrak{t}^{ij} + \mathfrak{m}^{ij}, \tag{4.16}$$

where

$$\mathfrak{m}^{ij} = \varepsilon_0\sqrt{\mathfrak{g}}\,(E^i E^j - \tfrac{1}{2}g^{ij}E^2) \tag{4.17}$$

is a type of Maxwell stress tensor, or electromagnetic stress tensor, which is symmetric and has the same form in all materials.

As shown in [*1*], a stored energy function Σ which is invariant under rigid motions must reduce to a function of the quantities

$$\begin{aligned} \Sigma &= \bar{\Sigma}(C_{AB}, \Pi^A, X^A), \\ C_{AB} &\equiv g_{ij}\, x^i_A\, x^j_B, \\ \Pi^A &\equiv |(\boldsymbol{x}/\boldsymbol{X})|\, X^A_i\, \mathfrak{P}^i/\varrho_0, \end{aligned} \tag{4.18}$$

each of which is an absolute scalar under general transformations of the spatial coordinates x^i. One can then show that the material stress tensor ${}_{\mathsf{M}}\mathrm{t}^{ij}$ is given by

$${}_{\mathsf{M}}\mathrm{t}^{ij} = 2\varrho \frac{\partial \bar{\Sigma}}{\partial C_{AB}}\, x^i_A\, x^j_B, \tag{4.19}$$

analogous to the formula of ordinary finite elasticity theory (*cf.* TRUESDELL [*7*, Eq. (39.2)]).

Following this deluge of formalism and definitions of stress tensors in elastic dielectrics, perhaps it is best that we restate our position. The fundamental entity from the point of view of mechanics is the Cauchy stress t^{ij} which is a certain function of the fields X^K, X^K_i, φ_i, $\mathfrak{P}^i$, and g_{ij} fixed by the functional form of $\mathfrak{W}(X^K, X^K_i, \mathfrak{P}^i, g_{ij})$. The field equations and boundary conditions are (4.11) and (4.12). These equations are obviously independent of any decomposition of the Cauchy stress into a mechanical and an electromagnetic component. Decompositions of the type (4.13) and (4.16) may be used as intuitive guides in the construction of admissible constitutive equations for $\mathfrak{W}$ but are otherwise irrelevant. A distinguishing feature of the decomposition (4.16) is that $\mathfrak{m}^{ij}$ is independent of $\mathfrak{P}$ and of the deformation and has the same form in every material. It need not vanish in a vacuum where $\mathfrak{W} = \mathfrak{P} = 0$. On the other hand, the material stress ${}_{\mathsf{M}}\mathrm{t}^{ij}$ is a function only of the polarization $\mathfrak{P}$ and the deformation and depends on these variables in a way which is characteristic of the material.

Consider next the Lagrange derivatives of $\mathscr{L}$ with respect to the field variables X^K, $\mathfrak{P}^i$, and φ. Let us denote them by

$$\begin{aligned} \mathfrak{L}_K &\equiv \frac{\partial \mathfrak{W}}{\partial X^K} - \partial_i\left(\frac{\partial \mathfrak{W}}{\partial X^K_i}\right), \qquad -\mathfrak{L}_i \equiv {}_{\mathsf{L}}E_i + E_i, \\ \mathfrak{L} &\equiv \varepsilon_0 \sqrt{\mathfrak{g}}\, \nabla^2 \varphi - \operatorname{div} \mathfrak{P}. \end{aligned} \tag{4.20}$$

For elastic dielectrics, TRAUTMAN's identity (2.42) takes the form

$$\nabla_i \mathrm{t}^i_j = \mathfrak{L}_K X^K_j - \mathfrak{L}_i\, \partial_j \mathfrak{P}^i - \mathfrak{L}\, E_j + (\delta^m_j \delta^i_k - \delta^i_j \delta^m_k)\, \partial_i(\mathfrak{L}_m \mathfrak{P}^k). \tag{4.21}$$

Since $\det X^K_j \neq 0$, it follows from (4.21) that the field equations (4.11) are equivalent to the set of Lagrange equations $\mathfrak{L}_K = 0$, $\mathfrak{L}_i = 0$, $\mathfrak{L} = 0$. That is, the field equation $\nabla_i \mathrm{t}^i_j = 0$ may be replaced by the equation $\mathfrak{L}_K = 0$ in the set (4.11).

Finally, we consider briefly the problem treated in [*1*] of an elastic dielectric placed in a vacuum. The variational principle of §3 still applies provided we take into account the constraint $\mathfrak{W} = 0$, $\mathfrak{P} = 0$ outside the dielectric. Here we let the discontinuity surface $\mathscr{S}(\lambda)$ correspond to the boundary between the dielectric and vacuum. The field equations interior to the dielectric retain the form (4.10), but owing to the constraints $\mathfrak{P} = \mathfrak{W} = 0$, exterior to the dielectric,

we get the field equations

$$\nabla_j \mathfrak{t}^j_k = 0, \qquad \varepsilon_0 \sqrt{\mathfrak{g}}\, \nabla^2 \varphi = 0. \tag{4.22}$$

But when $\mathfrak{W} = \mathfrak{P} = 0$, $\mathfrak{t}^i_j$ has the form

$$\mathfrak{t}^i_j = \varepsilon_0 \sqrt{\mathfrak{g}} (E^i E_j - \tfrac{1}{2} \delta^i_j E^2) \tag{4.23}$$

and Trautman's identity reduces to the trivial one

$$\nabla_i \mathfrak{t}^i_j = \varepsilon_0 \sqrt{\mathfrak{g}}\, \nabla^2 \varphi E_j. \tag{4.24}$$

Thus the field equations (4.22) are not independent, and we may take the simple scalar equation $(4.22)_2$ as the single vacuum field equation.

References

[1] Toupin, R. A.: The elastic dielectric. J. Rational Mech. Anal. **5**, 849–916 (1956).

[2] Minkowski, H.: Die Grundgleichungen für die elektromagnetischen Vorgänge in bewegten Körpern. Math. Ann. **68**, 472–525 (1910).

[3] Abraham, M.: Zur Elektrodynamik bewegter Körper. Rend. Palermo **28**, 1–28 (1909).

[4] Gyorgyi, G.: Die Bewegung des Energiemittelpunktes und der Energie-Impuls-Tensor des Elektro-Magnetischen Feldes in Dielektrika. Acta Phys. Hung. **4**, 121–131 (1954).

[5] Møller, C.: The Theory of Relativity. Oxford: Clarendon Press 1952.

[6] Schouten, J. A.: Ricci Calculus. Grundl. math. Wiss., 2. Aufl., Bd. 10. Berlin-Göttingen-Heidelberg: Springer 1954.

[7] Truesdell, C.: The mechanical foundations of elasticity and fluid mechanics. J. Rational Mech. Anal. **1**, 125–300 (1952); **2**, 593–616 (1953).

[8] Bergmann, P.: Non-linear field theories. Phys. Rev. **75**, 680–685 (1949).

[9] Thomas, T. Y.: Differential Invariants of Generalized Spaces. Cambridge 1934.

[10] Yano, K.: The Theory of Lie Derivatives and Its Applications. New York: Interscience Publishers Inc. 1957.

[11] Trautman, A.: On the conservation theorems and equations of motion in covariant field theories. Bull. Acad. Polon. Sci. **4**, 675–678 (1956).

[12] Weiss, P.: On the Hamilton-Jacobi theory and quantization of a dynamical continuum. Proc. Roy. Soc. **169**, 102–119 (1938).

[13] Einstein, A.: Hamiltonsches Prinzip und allgemeine Relativitätstheorie. S.-B. preuß. Akad. Wiss. 1111–1116 (1916).

U. S. Naval Research Laboratory
Washington D. C.

(Received March 2, 1960)

CONTENTS

Druck der Universitätsdruckerei H. Stürtz AG., Würzburg
Printed in Germany